



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification:</b> <b>G06F</b>	<b>A2</b>	<b>(11) International Publication Number:</b> <b>WO 00/33160</b> <b>(43) International Publication Date:</b> 08 June 2000 (08.06.2000)
<b>(21) International Application Number:</b> PCT/US99/28528 <b>(22) International Filing Date:</b> 02 December 1999 (02.12.1999) <b>(30) Priority Data:</b> 09/204,888 03 December 1998 (03.12.1998) US 09/205,119 03 December 1998 (03.12.1998) US 09/205,653 03 December 1998 (03.12.1998) US <b>(60) Parent Application or Grant</b> TELECOM PARTNERS LTD. [/]; (). ELDERING, Charles, A. [/]; (). SYLLA, M., Lamine [/]; (). ELDERING, Charles, A. [/]; (). SYLLA, M., Lamine [/]; (). BLASKO, John, P.; ().		<b>Published</b>
<b>(54) Title: SUBSCRIBER CHARACTERIZATION AND ADVERTISEMENT MONITORING SYSTEM</b> <b>(54) Titre: SYSTEME DE DETERMINATION D'UN PROFIL D'ABONNE ET DE SURVEILLANCE PUBLICITAIRE</b>  <b>(57) Abstract</b> <p>A subscriber characterization and advertisement monitoring system (100) is presented in which subscriber viewing habits are monitored to determine demographic profiles. These profiles can be utilized for the matching of advertisements to subscribers based on their viewing habits and estimated demographics and product interests. The system (100) can be run locally in a television set-top (1808) or can be run in client server mode where channel selections are transmitted from the residence (1800) to a centralized switching location (server) (1840) such as a telephone office or Internet Service Provider. In client-server mode the channel selections are monitored at the centralized location (1840) which also performs the subscriber characterization. The system also provides the ability to monitor if advertisements were viewed and for what duration.</p> <b>(57) Abrégé</b> <p>L'invention concerne un système de détermination de profil d'abonné et de surveillance publicitaire (100), dans lequel on contrôle les publicités qu'un abonné regarde, afin de déterminer des profils démographiques. On peut utiliser ces profils pour faire correspondre les publicités destinées à des abonnés en fonction de ce qu'ils regardent, l'évaluation des données démographiques, et l'intérêt présenté par un produit. Le système (100) peut être mis en application localement dans un appareil de télévision (1808), ou peut être utilisé en mode serveur-client, les sélections de chaînes étant transmises de la résidence (1800) vers l'emplacement de commutation (1840) centralisée (serveur), tel qu'un central téléphonique ou un fournisseur de services Internet. En mode serveur-client, les sélections de chaînes sont contrôlées au niveau de l'emplacement de commutation (1840) centralisé, ce mode déterminant également le profil d'abonné. Le système permet également de contrôler si les publicités sont regardées et pendant combien de temps.</p>		

PCT

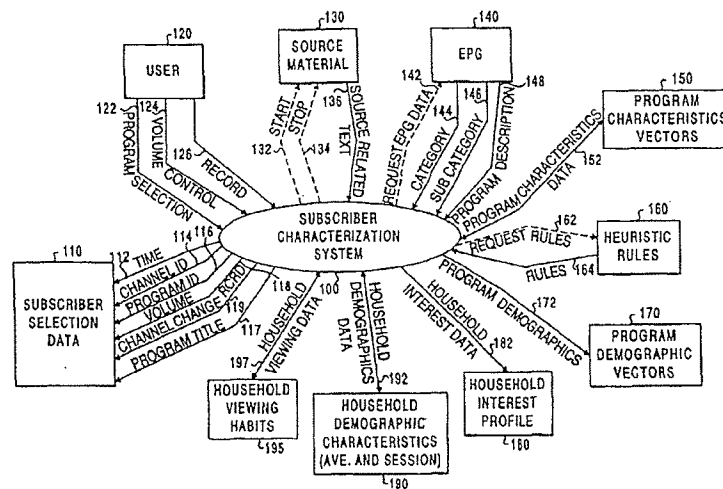
WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>7</sup> : <b>G06F</b>		A2	(11) International Publication Number: <b>WO 00/33160</b>
			(43) International Publication Date: 8 June 2000 (08.06.00)
(21) International Application Number: PCT/US99/28528		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 2 December 1999 (02.12.99)			
(30) Priority Data: 09/204,888 3 December 1998 (03.12.98) US 09/205,119 3 December 1998 (03.12.98) US 09/205,653 3 December 1998 (03.12.98) US			
(71) Applicant (for all designated States except US): TELECOM PARTNERS LTD. [US/US]; 900 Town Center, New Britain, PA 18901 (US).			
(72) Inventors; and (75) Inventors/Applicants (for US only): ELDERING, Charles, A. [US/US]; 315 Hedgerow Lane, Doylestown, PA 18901 (US). SYLLA, M., Lamine [SN/US]; 6 West Butler Avenue, New Britain, PA 18901 (US).			
(74) Agents: BLASKO, John, P. et al.; J.P. Blasko Professional Corp., 107 North Broad Street, Doylestown, PA 18901 (US).		Published Without international search report and to be republished upon receipt of that report.	

(54) Title: SUBSCRIBER CHARACTERIZATION AND ADVERTISEMENT MONITORING SYSTEM



(57) Abstract

A subscriber characterization and advertisement monitoring system (100) is presented in which subscriber viewing habits are monitored to determine demographic profiles. These profiles can be utilized for the matching of advertisements to subscribers based on their viewing habits and estimated demographics and product interests. The system (100) can be run locally in a television set-top (1808) or can be run in client server mode where channel selections are transmitted from the residence (1800) to a centralized switching location (server) (1840) such as a telephone office or Internet Service Provider. In client-server mode the channel selections are monitored at the centralized location (1840) which also performs the subscriber characterization. The system also provides the ability to monitor if advertisements were viewed and for what duration.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## Description

5

10

15

20

25

30

35

40

45

50

55

5

## TITLE

*Subscriber characterization and  
advertisement monitoring system*

10

5

## Background of the Invention

15

20

25

30

35

40

45

50

55

Cable television service providers have typically provided one-way broadcast services but now offer high-speed data services and can combine traditional analog broadcasts with digital broadcasts and access to Internet web sites. Telephone companies can offer digital data and video programming on a switched basis over digital subscriber line technology. Although the subscriber may only be presented with one channel at a time, channel change requests are instantaneously transmitted to centralized switching equipment and the subscriber can access the programming in a broadcast-like manner. Internet Service Providers (ISPs) offer Internet access and can offer access to text, audio, and video programming which can also be delivered in a broadcast-like manner in which the subscriber selects "channels" containing programming of interest. Such channels may be offered as part of a video programming service or within a data service and can be presented within an Internet browser.

Advertisements are a part of daily life and certainly an important part of entertainment programming, where the payments for advertisements cover the cost of network television. A method, which provides a flexible billing plan to cable network users based on the amount of advertisements viewed is described in U.S. Patent No. 5,532,735, which discloses a method of advertisement selection for interactive services. A user associated with an interactive TV is presented with a program and a set of advertisements.

5 The user can indicate the amount of advertisements in the  
set of advertisements he wants to view.

10 While advertisements are sometimes beneficial to  
subscribers and deliver desired information regarding  
5 specific products or services, consumers generally view  
10 advertising as a "necessary evil" for broadcast-type  
entertainment. For example, a method for obtaining  
information on advertised services or products is described  
15 in U.S. Patent No. 5,708,478, which discloses a computer  
20 system for enabling radio listeners and television watchers  
to obtain advertising information. The system includes steps  
of determining whether an incoming video or audio signal  
includes advertisement specific data of an advertiser and  
capturing and storing the advertiser specific data.

15 Manufacturers pay an extremely high price to present,  
25 in 30 seconds or less, an advertisement for their product,  
which they hope a consumer will watch. Unfortunately for the  
manufacturer, the consumer frequently uses that interval of  
time to check the programming being presented on the other  
30 channels, and may not watch any of the advertisement.  
Alternately, the consumer may mute the channel and ignore  
what the manufacturer has presented. In any case the  
probability that the consumer has watched the advertisement  
35 is quite low. It is not until millions of dollars have been  
25 spent on an advertising campaign that a manufacturer can  
determine that the ads have been effective. This is  
40 presently accomplished by monitoring sales of the product or  
TV programs or channels viewed by users as disclosed in  
various public documents. As an example, U.S. Patent No.  
30 4,546,382 discloses a television and market research data  
45 collection system and method. A data collection unit  
containing a memory stores data as to which of the plurality  
of TV modes are in use, which TV channel is being viewed as  
50 well as input from a suitable optical scanning device for

collecting information about user's product purchases.  
Another system described in U.S. Patent No. 4,258,386  
discloses a television audience measuring system. The system  
monitors and stores information representative of channel  
identification, the time at which the channel is selected  
and the time at which the selection of a channel is  
terminated. U.S. Patent No. 5,608,445 discloses also a  
method and device for data capture in television viewer  
research. Devices are attached to a video installation in  
order to determine to which channel a set is tuned.

With the advent of the Internet manufacturers and  
service providers have found ways to selectively insert  
their advertisements based on a subscribers requests for  
information. As an example, an individual who searches for  
"cars" on the Internet may see an advertisement for a  
particular type of car. Various internet-based advertising  
use this method. The product literature from IMGIS Inc., "Ad  
Force," printed from the World Wide Web site  
[http://www.starpt.com/core/ad\\_Target.html](http://www.starpt.com/core/ad_Target.html) on June 30, 1998  
discloses an ad targeting system. The system delivers ads to  
web sites visitors based on the content of the web page,  
time of day, day of the week, keyword, by the number of  
times a visitor sees an advertisement and by the order in  
which a series of advertisements are shown to a visitor.  
Nevertheless, unless the subscriber actually goes to the  
advertised web site, there is no way to determine if the  
advertisement has been watched. As the content on the  
Internet migrates to multimedia programming including audio  
and video, the costs for the advertising will increase, but  
unless the advertiser can be sure that a significant  
percentage of the message was watched or observed, the  
advertising is ineffective. Prior art products for  
generating reports of ad campaign are generally PC-centric  
as described in various product literature which include the

5 product literature from DoubleClick Inc., "DoubleClick:  
Reporting," printed from the World Wide Web (WWW) site  
http://www.doubleclick.net/dart/howi\_repo.htm on June 19,  
1998, which discloses the reporting capabilities of  
10 DoubleClick's Dynamic Advertising Reporting & Targeting  
(DART) product. The information in the reports includes  
daily impressions by advertisement type, average impression  
per day of week and by hour of day. The average response  
15 rate per user is also included in the reports. The product  
literature from Netgravity Inc. "AdServer 3," printed from  
the World Wide Web site http://www.netgravity.com/products/  
on July 9, 1998 discloses Netgravity's Adserver 3 product  
20 for online advertisement. The product generates reports  
including the profiles of visitors who viewed an ad and site  
15 traffic throughout the day, week, month and year.

25 The product literature from Media Metrix "Frequently  
Asked Questions", printed from the World Wide Web site  
http://www.mediametrix.com/interact\_mmfaq.htm on June 30,  
1998 discloses Media Metrix software, PC Meter, that runs in  
30 the background of a PC and monitors everything being done on  
that machine. It determines who is using the PC by age,  
income, gender and geographic region and tracks usage of  
software application, commercial online services and  
35 detailed page level viewing of the World Wide Web. The  
marketing literature from Matchlogic Inc., "Centralized Ad  
25 Management," printed from the World Wide Web site  
http://www.matchlogic.com/docs/services2.htm on July 1, 1998  
discloses Matchlogic services for ad management. The  
services include delivering advertisements based on pre-  
40 defined targeting criteria, generating reports on how many  
unique viewers saw which banner and how many times it was  
viewed. The product literature from Accipiter Inc.,  
Accipiter AdManager 2.0," printed from the World Wide Web  
45 site http://www.accipiter.com/products/ADManager/fab.html on  
50



5 July 9, 1998 discloses Accipiter's ad management system.  
After delivering an advertisement based on pre-defined  
criteria, the system can generate reports on an ad campaign.  
The reports include visitors' demographic data, number of  
10 5 impressions and clicks generated from the entire site and by  
each ad and advertiser.

15 In order to deliver more targeted programming and  
advertising to subscribers, it is necessary to understand  
their likes and dislikes to a greater extent than is  
20 10 presently done today. Systems which identify subscriber  
preferences based on their purchases and responses to  
questionnaires allow for the targeted marketing of  
25 literature in the mail, but do not in any sense allow for  
the rapid and precise delivery of programming and  
15 advertising which is known to have a high probability of  
acceptance to the subscriber. Other systems give users the  
25 possibility to chose their programming as described in U.S.  
Patent No. 5,223,924 which discloses a system and method for  
automatically correlating user preferences with a TV program  
30 20 information database. The system includes a processor that  
performs "free text" search techniques to correlate the  
downloaded TV program information with the viewer's  
35 preferences. This system requires an interaction between the  
users and the programming. The white paper from Net  
25 Perceptions corporation entitled "Adding Value in the  
Digital Age" and printed from the World Wide Web site  
40 <http://www.netperceptions.com/products/white-papers.html> on  
June 30, 1998 discloses how the GroupLens Recommendation  
Engine gives online businesses the ability to target and  
30 30 personalize services, content, products and advertising. A  
45 learning process learns personal information about an  
individual using explicit and implicit ratings, a prediction  
process predicts user preference using collaborative

5 filtering and the recommendation process recommends products  
or services to users based on predictions.

10 The product literature from Aptex software Inc.,  
"SelectCast for Commerce Servers," printed from the World  
5 Wide Web site <http://www.aptex.com/products-selectcast-commerce.htm> on June 30, 1998 describes the product  
10 SelectCast for Commerce Servers. It personalizes online  
shopping based on observed user behavior. User interests are  
learned based on the content they browse, the promotions  
they click and the products they purchase.

15 In order to determine which programming or advertising  
is appropriate for the subscriber, knowledge of that  
subscriber and the subscriber product and programming  
20 preferences is required. Different methods are being used to  
gain knowledge of user's preferences and to profile the  
25 users. Generally, these methods use content or data mining  
technologies to profile users or predict their preferences.  
Another technique for predicting user's preferences is based  
on the use of collaborative filtering as described in U.S.  
30 Patent No. 5,704,017 which discloses a collaborative  
filtering system utilizing a belief network. The system  
learns a belief network using prior knowledge obtained from  
an expert in a given field of decision making and a database  
35 containing empirical data such as users' attributes as well  
as their preferences in that decision making field. The  
25 belief network can determine the probability of the unknown  
preferences of the user given the known attributes and thus  
40 predicts the preference most likely to be desired by the  
user.

30 The product literature from Aptex software Inc.,  
45 "SelectCast for Ad Servers," printed from the World Wide Web  
site <http://www.aptex.com/products-selectcast-ads.htm> on  
June 30, 1998 discloses an ad targeting system from Aptex  
Software Inc. The system employs neural networks and a

5 context vector data model to optimize relationships between  
users and content. It provides user profiling by mining the  
context and content of all actions including clicks,  
queries, page views and ad impressions. Aptex's technology  
10 5 uses a context vector data modeling technique described in  
U.S. Patent No. 5,619,709 which discloses a system and  
method of context vector generation and retrieval. Context  
vectors represent conceptual relationships among information  
15 items by quantitative means. A neural network operates on a  
training corpus of records to develop relationship-based  
context vectors based on word proximity and co-importance.  
Geometric relationships among context vectors are  
20 representative of conceptual relationships among their  
associated items.

15 The product Data sheet from Open Sesame, "Learn  
Sesame," printed from the World Wide Web site  
25 [http://www.opensesame.com/prod\\_04.html](http://www.opensesame.com/prod_04.html) on July 09, 1998  
discloses Open Sesame's personalization product for Web  
enterprises. It learns about users automatically from their  
30 20 browsing behavior.

The product literature from Engage Technologies,  
"Engage.Discover," printed from the World Wide Web site  
35 <http://www.engagetech.com> on July 09, 1998 discloses Engage  
Technologies' product for user profiling. User-disclosed  
25 information such as interest, demographics and opinions are  
combined with anonymous clickstream data that describes  
40 where users come from before visiting the site, how long  
they stay, and what pages or types of pages they visit most  
frequently to build the visitor profile.

30 The marketing literature from Broadvision, "The Power  
of Personalization", printed from the World Wide Web site  
45 [http://www.broadvision.com/content/corporate/brochure/Broch4](http://www.broadvision.com/content/corporate/brochure/Broch4.htm)  
.htm on August 21, 1998 discloses BroadVision One-to-One  
50 application profiling system. The system learns about users

5 through a variety of techniques including registration,  
questionnaires, observation and integration of historical  
and externally generated data.

10 The marketing literature from Firefly Corporation,  
5 "Firefly passport Office," printed from the World Wide Web  
site <http://www.firefly.net/company/PassportOffice.html> on  
June 20, 1998 discloses Firefly's Relationship Management  
software. The software enables online businesses to create,  
15 extend and manage personal profiles for every user.

20 Specific information regarding a subscriber's viewing  
habits or the Internet web sites they have accessed can be  
stored for analysis, but such records are considered private  
and subscribers are not generally willing to have such  
information leave their control. Although there are  
15 regulatory models, which permit the collection of such data  
on a "notice and consent" basis, there is a general tendency  
towards legal rules, which prohibit such raw data to be  
collected.

30 With the migration of services from a broadcast based  
20 model to a client-server based model in which subscribers  
make individualized request for programming to an Internet  
access provider or content provider, there is opportunity to  
monitor the subscriber viewing characteristics to better  
35 provide them with programming and advertising which will be  
of interest to them. A server may act as a proxy for the  
subscriber requests and thus be able to monitor what a  
subscriber has requested and is viewing. Since subscribers  
40 may not want this raw data to be utilized, there is a need  
for a system which can process this information and generate  
30 statistically relevant subscriber profiles. These profiles  
45 should be accessible to others on the network who may wish  
to determine if their programming or advertisements are  
suitable for the subscriber. In a broadcast-based model, the  
information to be processed can be embedded within the TV

5 program or broadcast separately and can be in form of an  
electronic program guide (EPG) or text information related  
to the program. As an example, U.S. Patent No. 5,579,055  
discloses an electronic program guide (EPG) and text channel  
10 5 data controller. The text and EPG data are embedded in the  
vertical blanking interval of the video signal and  
extracted, at reception, by the data controller. The EPG  
contains information fields such as program category,  
15 program subcategory and program content description. U.S.  
10 Patent No. 5,596,373 discloses also a method and apparatus  
for providing program oriented information in a multiple  
station broadcasting system. The EPG data includes guide  
20 data, channel data and program data. The program data  
includes among other information, the program title, the  
15 program category, the program sub-category and a detailed  
description of the program.

For the foregoing reasons, there is a need for an  
advertisement monitoring system which can monitor which  
advertisements have been viewed by a subscriber. There is  
30 20 also a need for a subscriber characterization system which  
can generate and store subscriber characteristics which  
reflect the probable demographics and preferences of the  
subscriber and household.

25

#### Summary Of The Invention

40 The present invention encompasses a system for  
determining to what extent an advertisement has been viewed  
by a subscriber or household.

45 30 In a preferred embodiment subscriber selection data  
including the channel selected and the time at which it was  
selected are recorded. Advertisement related information  
including the type of product, brand name, and other  
50 descriptive information which categorizes the advertisement

5 is extracted from the advertisement or text information  
related to the advertisement including closed captioning  
text. Based on the subscriber selection data a record of  
what percentage of the advertisement was watched is created.

10 5 This record can subsequently be used to make a measure of  
the effectiveness of the advertisement.

15 In a preferred embodiment the text information related  
to the advertisement is processed using context mining  
techniques which allow for classification of the  
10 advertisement and extraction of key data including product  
type and brand. Context mining techniques allow for  
determination of a product type, product brand name and in  
20 the case of a product which is not sold with a particular  
brand name, a generic name for the product.

15 The present invention can also be realized in a client-  
25 server mode in which case the subscriber executes channel  
changes at the client side of the network which are  
transmitted to the server side and fulfilled by the routing  
of a channel to the subscriber. The server side monitors  
30 the subscriber activity and stores the record of channel  
change requests. Advertisement related information is  
retrieved from the server side, which contains the  
advertising material itself, retrieves the advertising  
35 material from a third party, or analyzes the data stream  
25 carrying the advertising to the subscriber. The server side  
extracts descriptive fields from the advertisement and based  
40 on the subscriber selection data, determines the extent to  
which the advertisement was viewed by the subscriber. As an  
example the system can determine the percentage of the  
30 advertisement that was viewed by the subscriber.

45 The present invention includes a system for  
characterizing subscribers watching video or multimedia  
programming based on monitoring their detailed selection  
50 choices including the time duration of their viewing, the

5 volume the programming is listened at, the program  
selection, and collecting text information about that  
programming to determine what type of programming the  
subscriber is most interested in. In addition, the system  
10 5 can generate a demographic description of the subscriber or  
household which describes the probable age, income, gender  
and other demographics. The resulting characterization  
includes probabilistic determinations of what other  
15 programming or products the subscriber/household will be  
10 interested in.

In a preferred embodiment, the textual information  
which describes the programming is obtained by context  
20 mining of text associated with the programming. The  
associated text can be from the closed-captioning data  
15 associated with the programming, an electronic program  
guide, or from text files associated with or part of the  
25 programming itself.

The system can provide both session measurements which  
correspond to a profile obtained over a viewing session, or  
30 20 an average profile which corresponds to data obtained over  
multiple viewing sessions.

The present invention also encompasses the use of  
heuristic rules in logical form or expressed as conditional  
35 probabilities to aid in forming a subscriber profile. The  
25 heuristic rules in logical form allow the system to apply  
generalizations which have been learned from external  
40 studies to obtain a characterization of the subscriber. In  
the case of conditional probabilities, determinations of the  
probable content of a program can be applied in a  
30 mathematical step to a matrix of conditional probabilities  
45 to obtain probabilistic subscriber profiles indicating  
program and product likes and dislikes as well for  
determining probabilistic demographic data.

50 One advantage of the present invention is that it

5 allows consumers the possibility to permit access to  
probabilistic information regarding their household  
demographics and programming/product preferences, without  
revealing their specific viewing history. Subscribers may  
10 elect to permit access to this information in order to  
receive advertising which is more targeted to their  
likes/dislikes. Similarly, a subscriber may wish to sell  
access to this statistical data in order to receive revenue  
15 or receive a discount on a product or a service.

10 Another advantage of the present invention is that the  
resulting probabilistic information can be stored locally  
and controlled by the subscriber, or can be transferred to a  
20 third party which can provide access to the subscriber  
characterization. The information can also be encrypted to  
15 prevent unauthorized access in which case only the  
subscriber or someone authorized by the subscriber can  
25 access the data.

The present invention includes also a system for  
characterizing subscribers watching video or multimedia  
30 programming based on monitoring the requests made by the  
subscriber for programming to a server which contains the  
content or which requests the content from a third party.  
35 The server side of the network is able to monitor the  
subscriber's detailed selection choices including the time  
25 duration of their viewing, the volume the programming is  
listened at, and the program selection.

40 The server side collects text information about that  
programming to determine what type of programming the  
subscriber is most interested in. In addition the system  
30 can generate a demographic description of the subscriber or  
45 household which describes the probable age, income, gender  
and other demographics. The resulting characterization  
includes probabilistic determinations of what other



programming or products the subscriber/household will be  
interested in.

These and other features and objects of the invention  
will be more fully understood from the following detailed  
description of the preferred embodiments which should be  
read in light of the accompanying drawings.

#### Brief Description of the Drawings

The accompanying drawings, which are incorporated in  
and form a part of the specification, illustrate the  
embodiments of the present invention and, together with the  
description serve to explain the principles of the  
invention.

In the drawings:

FIG. 1 shows a context diagram for a subscriber  
characterization system.

FIG. 2 illustrates a block diagram for a realization of  
a subscriber monitoring system for receiving video signals;

FIG. 3 illustrates a block diagram of a channel  
processor;

FIG. 4 illustrates a block diagram of a computer for a  
realization of the subscriber monitoring system;

FIG. 5 illustrates a channel sequence and volume over a  
twenty-four (24) hour period;

FIG. 6 illustrates a time of day detailed record;

FIG. 7 illustrates a household viewing habits  
statistical table;

FIG. 8A illustrates an entity-relationship diagram for  
the generation of program characteristics vectors;

FIG. 8B illustrates a flowchart for program  
characterization;

FIGS. 9A illustrates a deterministic program category  
vector;

FIG. 9B illustrates a deterministic program sub-  
category vector;

FIG. 9C illustrates a deterministic program rating  
vector;

FIG. 9D illustrates a probabilistic program category  
vector;

FIG. 9E illustrates a probabilistic program sub-  
category vector;

FIG. 9F illustrates a probabilistic program content  
vector;

FIG. 10A illustrates a set of logical heuristic rules;

FIG. 10B illustrates a set of heuristic rules expressed  
in terms of conditional probabilities;

FIG. 11 illustrates an entity-relationship diagram for  
the generation of program demographic vectors;

FIG. 12 illustrates a program demographic vector;

FIG. 13 illustrates an entity-relationship diagram for  
the generation of household session demographic data and  
household session interest profiles;

FIG. 14 illustrates an entity-relationship diagram for  
the generation of average and session household demographic  
characteristics;

FIG. 15 illustrates average and session household  
demographic data;

FIG. 16 illustrates an entity-relationship diagram for  
generation of a household interest profile;

FIG. 17 illustrates household interest profile  
including programming and product profiles;

FIG. 18 illustrates a client-server architecture for  
realizing the present invention; and

FIG. 19 illustrates an advertisement monitoring table.

## Detailed Description

## Of The Preferred Embodiment

In describing a preferred embodiment of the invention illustrated in the drawings, specific terminology will be used for the sake of clarity. However, the invention is not intended to be limited to the specific terms so selected, and it is to be understood that each specific term includes all technical equivalents which operate in a similar manner to accomplish a similar purpose.

With reference to the drawings, in general, and FIGS. 1 through 19 in particular, the apparatus of the present invention is disclosed.

The present invention is directed at an apparatus for monitoring which advertisements are watched by a subscriber or a household.

In the present system the programming viewed by the subscriber, both entertainment and advertisement, can be studied and processed by the subscriber characterization system to determine the program characteristics. This determination of the program characteristics is referred to as a program characteristics vector. The vector may be a truly one-dimensional vector, but can also be represented as an  $n$  dimensional matrix which can be decomposed into vectors. For advertisements, the program characteristics vector can contain information regarding the advertisement including product type, features, brand or generic name, or other relevant advertising information.

The subscriber profile vector represents a profile of the subscriber (or the household of subscribers) and can be in the form of a demographic profile (average or session) or a program or product preference vector. The program and product preference vectors are considered to be part of a household interest profile which can be thought of as an  $n$  dimensional matrix representing probabilistic measurements

of subscriber interests.

5 In the case that the subscriber profile vector is a  
demographic profile, the subscriber profile vector indicates  
a probabilistic measure of the age of the subscriber or  
10 5 average age of the viewers in the household, sex of the  
subscriber, income range of the subscriber or household, and  
other such demographic data. Such information comprises  
household demographic characteristics and is composed of  
15 both average and session values. Extracting a single set of  
10 values from the household demographic characteristics can  
correspond to a subscriber profile vector.

20 The household interest profile can contain both  
programming and product profiles, with programming profiles  
corresponding to probabilistic determinations of what  
15 programming the subscriber (household) is likely to be  
interested in, and product profiles corresponding to what  
25 products the subscriber (household) is likely to be  
interested in. These profiles contain both an average value  
and a session value, the average value being a time average  
30 20 of data, where the averaging period may be several days,  
weeks, months, or the time between resets of unit.

Since a viewing session is likely to be dominated by a  
particular viewer, the session values may, in some  
35 circumstances, correspond most closely to the subscriber  
values, while the average values may, in some circumstances,  
25 correspond most closely to the household values.

40 FIG. 1 depicts the context diagram of a preferred  
embodiment of a Subscriber Characterization System (SCS)  
100. A context diagram, in combination with entity-  
30 relationship diagrams, provide a basis from which one  
45 skilled in the art can realize the present invention. The  
present invention can be realized in a number of programming  
languages including C, C++, Perl, and Java, although the  
scope of the invention is not limited by the choice of a  
50

5 particular programming language or tool. Object oriented  
languages have several advantages in terms of construction  
of the software used to realize the present invention,  
although the present invention can be realized in procedural  
10 5 or other types of programming languages known to those  
skilled in the art.

15 In generating a subscriber profile, the SCS 100  
receives from a user 120 commands in the form of a volume  
control signal 124 or program selection data 122 which can  
20 10 be in the form of a channel change but may also be an  
address request which requests the delivery of programming  
from a network address. A record signal 126 indicates that  
the programming or the address of the programming is being  
recorded by the user. The record signal 126 can also be a  
25 15 printing command, a tape recording command, a bookmark  
command or any other command intended to store the program  
being viewed, or program address, for later use.

30 The material being viewed by the user 120 is referred  
to as source material 130. The source material 130, as  
35 20 defined herein, is the content that a subscriber selects and  
may consist of analog video, Motion Picture Expert Group  
(MPEG) digital video source material, other digital or  
analog material, Hypertext Markup Language (HTML) or other  
type of multimedia source material. The subscriber  
40 25 characterization system 100 can access the source material  
130 received by the user 120 using a start signal 132 and a  
stop signal 134, which control the transfer of source  
related text 136 which can be analyzed as described herein.

45 In a preferred embodiment, the source related text 136  
30 can be extracted from the source material 130 and stored in  
memory. The source related text 136, as defined herein,  
includes source related textual information including  
descriptive fields which are related to the source material  
50 130, or text which is part of the source material 130

5 itself. The source related text 136 can be derived from a  
number of sources including but not limited to closed  
captioning information, Electronic Program Guide (EPG)  
material, and text information in the source itself (e.g.  
5 text in HTML files).

10 Electronic Program Guide (EPG) 140 contains information  
related to the source material 130 which is useful to the  
user 120. The EPG 140 is typically a navigational tool  
15 which contains source related information including but not  
limited to the programming category, program description,  
rating, actors, and duration. The structure and content of  
EPG data is described in detail in US Patent 5,596,373  
20 assigned to Sony Corporation and Sony Electronics which is  
herein incorporated by reference. As shown in FIG. 1, the  
15 EPG 140 can be accessed by the SCS 100 by a request EPG data  
signal 142 which results in the return of a category 144, a  
25 sub-category 146, and a program description 148. EPG  
information can potentially include fields related to  
advertising.

30 20 In one embodiment of the present invention, EPG data is  
accessed and program information such as the category 144,  
the sub-category 146, and the program description 148 are  
stored in memory.

35 In another embodiment of the present invention, the  
25 source related text 136 is the closed captioning text  
embedded in the analog or digital video signal. Such closed  
captioning text can be stored in memory for processing to  
40 extract the program characteristic vectors 150.

One of the functions of the SCS 100 is to generate the  
30 program characteristics vectors 150 which are comprised of  
program characteristics data 152, as illustrated in FIG. 1.  
45 The program characteristics data 152, which can be used to  
create the program characteristics vectors 150 both in  
vector and table form, are examples of source related  
50

5 information which represent characteristics of the source  
material. In a preferred embodiment, the program  
characteristics vectors 150 are lists of values which  
10 characterize the programming (source) material in according  
to the category 144, the sub-category 146, and the program  
description 148. The present invention may also be applied  
to advertisements, in which case program characteristics  
15 vectors contain, as an example, a product category, a  
product sub-category, and a brand name.

10 As illustrated in FIG. 1, the SCS 100 uses heuristic  
rules 160. The heuristic rules 160, as described herein,  
are composed of both logical heuristic rules as well as  
20 heuristic rules expressed in terms of conditional  
probabilities. The heuristic rules 160 can be accessed by  
15 the SCS 100 via a request rules signal 162 which results in  
the transfer of a copy of rules 164 to the SCS 100.

25 The SCS 100 forms program demographic vectors 170 from  
program demographics 172, as illustrated in FIG. 1. The  
program demographic vectors 170 also represent  
30 20 characteristics of source related information in the form of  
the intended or expected demographics of the audience for  
which the source material is intended.

35 Subscriber selection data 110 is obtained from the  
monitored activities of the user and in a preferred  
25 embodiment can be stored in a dedicated memory. In an  
alternate embodiment, the subscriber selection data 110 is  
40 stored in a storage disk. Information which is utilized to  
form the subscriber selection data 110 includes time 112,  
which corresponds to the time of an event, channel ID 114,  
30 program ID 116, volume level 118, channel change record 119,  
45 and program title 117. A detailed record of selection data  
is illustrated in FIG. 6.

In a preferred embodiment, a household viewing habits  
50 195 illustrated in FIG. 1 is computed from the subscriber

5 selection data 110. The SCS 100 transfers household viewing  
data 197 to form household viewing habits 195. The  
household viewing data 197 is derived from the subscriber  
10 selection data 110 by looking at viewing habits at a  
5 particular time of day over an extended period of time,  
usually several days or weeks, and making some  
generalizations regarding the viewing habits during that  
time period.

15 The program characteristics vector 150 is derived from  
10 the source related text 136 and/or from the EPG 140 by  
applying information retrieval techniques. The details of  
this process are discussed in accordance with FIG. 8.

20 The program characteristics vector 150 is used in  
combination with a set of the heuristic rules 160 to define  
15 a set of the program demographic vectors 170 illustrated in  
FIG. 1 describing the audience the program is intended for.

25 One output of the SCS 100 is a household profile  
including household demographic characteristics 190 and a  
household interest profile 180. The household demographic  
30 characteristics 190 resulting from the transfer of household  
demographic data 192, and the household interest profile  
180, resulting from the transfer of household interests data  
182. Both the household demographics characteristics 190  
35 and the household interest profile 180 have a session value  
and an average value, as will be discussed herein.

40 The monitoring system depicted in FIG. 2 is responsible  
for monitoring the subscriber activities, and can be used to  
realize the SCS 100. In a preferred embodiment, the  
monitoring system of FIG. 2 is located in a television set-  
30 top device or in the television itself. In an alternate  
45 embodiment, the monitoring system is part of a computer  
which receives programming from a network.

In an application of the system for television  
50 services, an input connector 220 accepts the video signal



5 coming either from an antenna, cable television input, or  
other network. The video signal can be analog or Digital  
MPEG. Alternatively, the video source may be a video stream  
or other multimedia stream from a communications network  
10 including the Internet.

10 In the case of either analog or digital video, selected  
fields are defined to carry EPG data or closed captioning  
text. For analog video, the closed captioning text is  
15 embedded in the vertical blanking interval (VBI). As  
described in US Patent 5,579,005, assigned to Scientific-  
Atlanta, Inc., the EPG information can be carried in a  
dedicated channel or embedded in the VBI. For digital  
20 video, the closed captioning text is carried as video user  
bits in a user\_data field. The EPG data is transmitted as  
15 ancillary data and is multiplexed at the transport layer  
with the audio and video data.

25 Referring to FIG. 2, a system control unit 200 receives  
commands from the user 120, decodes the command and forwards  
the command to the destined module. In a preferred  
30 embodiment, the commands are entered via a remote control to  
a remote receiver 205 or a set of selection buttons 207  
available at the front panel of the system control unit 200.  
In an alternate embodiment, the commands are entered by the  
35 user 120 via a keyboard.

25 The system control unit 200 also contains a Central  
Processing Unit (CPU) 203 for processing and supervising all  
of the operations of the system control unit 200, a Read  
40 Only Memory (ROM) 202 containing the software and fixed  
data, a Random Access Memory (RAM) 204 for storing data. CPU  
30 203, RAM 204, ROM 202, and I/O controller 201 are attached  
to a master bus 206. A power supply in a form of battery  
45 can also be included in the system control unit 200 for  
backup in case of power outage.

50 An input/output (I/O) controller 201 interfaces the

5 system control unit 200 with external devices. In a preferred embodiment, the I/O controller 201 interfaces to the remote receiver 205 and a selection button such as the channel change button on a remote control. In an alternate  
10 5 embodiment, it can accept input from a keyboard or a mouse.

10 The program selection data 122 is forwarded to a channel processor 210. The channel processor 210 tunes to a selected channel and the media stream is decomposed into its  
15 basic components: the video stream, the audio stream, and the data stream. The video stream is directed to a video  
20 processor module 230 where it is decoded and further processed for display to the TV screen. The audio stream is directed to an audio processor 240 for decoding and output to the speakers.

15 The data stream can be EPG data, closed captioning text, Extended Data Service (EDS) information, a combination of these, or an alternate type of data. In the case of EDS the call sign, program name and other useful data are  
25 provided. In a preferred embodiment, the data stream is stored in a reserved location of the RAM 204. In an alternate embodiment, a magnetic disk is used for data storage. The system control unit 200 writes also in a  
30 dedicated memory, which in a preferred embodiment is the RAM 204, the selected channel, the time 112 of selection, the volume level 118 and the program ID 116 and the program  
35 title 117. Upon receiving the program selection data 122, the new selected channel is directed to the channel processor 210 and the system control unit 200 writes to the dedicated memory the channel selection end time and the  
40 program title 117 at the time 112 of channel change. The system control unit 200 keeps track of the number of channel changes occurring during the viewing time via the channel change record 119. This data forms part of the subscriber  
45 selection data 110.

5 The volume control signal 124 is sent to the audio  
processor 240. In a preferred embodiment, the volume level  
118 selected by the user 120 corresponds to the listening  
10 volume. In an alternate embodiment, the volume level 118  
5 selected by the user 120 represents a volume level to  
another piece of equipment such as an audio system (home  
theatre system) or to the television itself. In such a  
case, the volume can be measured directly by a microphone or  
15 other audio sensing device which can monitor the volume at  
20 which the selected source material is being listened.

A program change occurring while watching a selected  
channel is also logged by the system control unit 200.  
20 Monitoring the content of the program at the time of the  
program change can be done by reading the content of the  
15 EDS. The EDS contains information such as program title,  
which is transmitted via the VBI. A change on the program  
25 title field is detected by the monitoring system and logged  
as an event. In an alternate embodiment, an EPG is present  
and program information can be extracted from the EPG. In a  
30 20 preferred embodiment, the programming data received from the  
EDS or EPG permits distinguishing between entertainment  
programming and advertisements.

FIG. 3 shows the block diagram of the channel processor  
35 210. In a preferred embodiment, the input connector 220  
25 connects to a tuner 300 which tunes to the selected channel.  
A local oscillator can be used to heterodyne the signal to  
the IF signal. A demodulator 302 demodulates the received  
40 signal and the output is fed to an FEC decoder 304. The data  
stream received from the FEC decoder 304 is, in a preferred  
30 embodiment, in an MPEG format. In a preferred embodiment,  
45 system demultiplexer 306 separates out video and audio  
information for subsequent decompression and processing, as  
well as ancillary data which can contain program related  
information.

5 The data stream presented to the system demultiplexer  
306 consists of packets of data including video, audio and  
ancillary data. The system demultiplexer 306 identifies each  
packet from the stream ID and directs the stream to the  
10 corresponding processor. The video data is directed to the  
video processor module 230 and the audio data is directed to  
the audio processor 240. The ancillary data can contain  
closed captioning text, emergency messages, program guide,  
15 or other useful information.

20 Closed captioning text is considered to be ancillary  
data and is thus contained in the video stream. The system  
demultiplexer 306 accesses the user data field of the video  
stream to extract the closed captioning text. The program  
guide, if present, is carried on data stream identified by a  
25 specific transport program identifier.

30 In an alternate embodiment, analog video can be used.  
For analog programming, ancillary data such as closed  
captioning text or EDS data are carried in a vertical  
blanking interval.

35 FIG. 4 shows the block diagram of a computer system for  
a realization of the subscriber monitoring system based on  
the reception of multimedia signals from a bi-directional  
network. A system bus 422 transports data amongst the CPU  
203, the RAM 204, Read Only Memory - Basic Input Output  
25 System (ROM-BIOS) 406 and other components. The CPU 203  
accesses a hard drive 400 through a disk controller 402. The  
standard input/output devices are connected to the system  
bus 422 through the I/O controller 201. A keyboard is  
40 attached to the I/O controller 201 through a keyboard port  
416 and the monitor is connected through a monitor port 418.  
30 The serial port device uses a serial port 420 to communicate  
with the I/O controller 201. Industry Standard Architecture  
(ISA) expansion slots 408 and Peripheral Component  
Interconnect (PCI) expansion slots 410 allow additional  
50

cards to be placed into the computer. In a preferred embodiment, a network card is available to interface a local area, wide area, or other network.

FIG. 5 illustrates a channel sequence and volume over a twenty-four (24) hour period. The Y-axis represents the status of the receiver in terms of on/off status and volume level. The X-axis represents the time of day. The channels viewed are represented by the windows 501-506, with a first channel 502 being watched followed by the viewing of a second channel 504, and a third channel 506 in the morning. In the evening a fourth channel 501 is watched, a fifth channel 503, and a sixth channel 505. A channel change is illustrated by a momentary transition to the "off" status and a volume change is represented by a change of level on the Y-axis.

A detailed record of the subscriber selection data 110 is illustrated in FIG. 6 in a table format. A time column 602 contains the starting time of every event occurring during the viewing time. A Channel ID column 604 lists the channels viewed or visited during that period. A program title column 603 contains the titles of all programs viewed. A volume column 601 contains the volume level 118 at the time 112 of viewing a selected channel.

A representative statistical record corresponding to the household viewing habits 195 is illustrated in FIG. 7. In a preferred embodiment, a time of day column 700 is organized in period of time including morning, mid-day, afternoon, night, and late night. In an alternate embodiment, smaller time periods are used. A minutes watched column 702 lists, for each period of time, the time in minutes in which the SCS 100 recorded delivery of programming. The number of channel changes during that period and the average volume are also included in that table in a channel changes column 704 and an average volume

5 column 706 respectively. The last row of the statistical  
record contains the totals for the items listed in the  
minutes watched column 702, the channel changes column 704  
and the average volume 706.

10 5 FIG. 8A illustrates an entity-relationship diagram for  
the generation of the program characteristics vector 150.  
The context vector generation and retrieval technique  
described in US Patent 5,619,709, which is incorporated  
15 herein by reference, can be applied for the generation of  
10 the program characteristics vectors 150. Other techniques  
are well known by those skilled in the art.

20 Referring to FIG. 8A, the source material 130 or the  
EPG 140 are passed through a program characterization  
process 800 to generate the program characteristics vectors  
15 150. The program characterization process 800 is described  
25 in accordance with FIG. 8B. Program content descriptors  
including a first program content descriptor 802, a second  
program content descriptor 804 and an nth program content  
descriptor 806, each classified in terms of the category  
30 144, the sub-category 146, and other divisions as identified  
in the industry accepted program classification system, are  
presented to a context vector generator 820. As an example,  
35 the program content descriptor can be text representative of  
the expected content of material found in the particular  
25 program category 144. In this example, the program content  
descriptors 802, 804 and 806 would contain text  
representative of what would be found in programs in the  
40 news, fiction, and advertising categories respectively. The  
context vector generator 820 generates context vectors for  
30 that set of sample texts resulting in a first summary  
context vector 808, a second summary context vector 810, and  
45 an nth summary context vector 812. In the example given, the  
summary context vectors 808, 810, and 812 correspond to the  
categories of news, fiction and advertising respectively.

5 The summary vectors are stored in a local data storage system.

10 Referring to FIG. 8B, a sample of the source related text 136 which is associated with the new program to be classified is passed to the context vector generator 820 which generates a program context vector 840 for that program. The source related text 136 can be either the source material 130, the EPG 140, or other text associated with the source material. A comparison is made between the actual program context vectors and the stored program content context vectors by computing, in a dot product computation process 830, the dot product of the first summary context vector 808 with the program context vector 840 to produce a first dot product 814. Similar operations are performed to produce second dot product 816 and nth dot product 818.

25 The values contained in the dot products 814, 816 and 818, while not probabilistic in nature, can be expressed in probabilistic terms using a simple transformation in which the result represents a confidence level of assigning the corresponding content to that program. The transformed values add up to one. The dot products can be used to classify a program, or form a weighted sum of classifications which results in the program characteristics vectors 150. In the example given, if the source related text 136 was from an advertisement, the nth dot product 818 would have a high value, indicating that the advertising category was the most appropriate category, and assigning a high probability value to that category. If the dot products corresponding to the other categories were significantly higher than zero, those categories would be assigned a value, with the result being the program characteristics vectors 150 as shown in FIG. 9D.

50 For the sub-categories, probabilities obtained from the

5 content pertaining to the same sub-category 146 are summed  
to form the probability for the new program being in that  
sub-category 146. At the sub-category level, the same method  
10 is applied to compute the probability of a program being  
5 from the given category 144. The three levels of the program  
classification system; the category 144, the sub-category  
146 and the content, are used by the program  
characterization process 800 to form the program  
15 characteristics vectors 150 which are depicted in FIGS. 9D-  
10 9F.

The program characteristics vectors 150 in general are  
represented in FIGS. 9A through 9F. FIGS. 9A, 9B and 9C are  
20 an example of deterministic program vectors. This set of  
vectors is generated when the program characteristics are  
15 well defined, as can occur when the source related text 136  
or the EPG 140 contains specific fields identifying the  
category 144 and the sub-category 146. A program rating can  
25 also be provided by the EPG 140.

In the case that these characteristics are not  
30 specified, a statistical set of vectors is generated from  
the process described in accordance with FIG. 8. FIG. 9D  
shows the probability that a program being watched is from  
the given category 144. The categories are listed in the X-  
35 axis. The sub-category 146 is also expressed in terms of  
probability. This is shown in FIG. 9E. The content component  
25 of this set of vectors is a third possible level of the  
program classification, and is illustrated in FIG. 9F.

FIG. 10A illustrates sets of logical heuristic rules  
which form part of the heuristic rules 160. In a preferred  
30 embodiment, logical heuristic rules are obtained from  
sociological or psychological studies. Two types of rules  
45 are illustrated in FIG. 10A. The first type links an  
individual's viewing characteristics to demographic  
characteristics such as gender, age, and income level. A  
50



5 channel changing rate rule 1030 attempts to determine gender  
based on channel change rate. An income related channel  
change rate rule 1010 attempts to link channel change rates  
to income brackets. A second type of rules links particular  
10 5 programs to particular audience, as illustrated by a gender  
determining rule 1050 which links the program category  
144/sub-category 146 with a gender. The result of the  
application of the logical heuristic rules illustrated in  
15 FIG. 10A are probabilistic determinations of factors  
including gender, age, and income level. Although a specific  
set of logical heuristic rules has been used as an example,  
20 a wide number of types of logical heuristic rules can be  
used to realize the present invention. In addition, these  
rules can be changed based on learning within the system or  
15 based on external studies which provide more accurate rules.

25 FIG. 10B illustrates a set of the heuristic rules 160  
expressed in terms of conditional probabilities. In the  
example shown in FIG. 10B, the category 144 has associated  
with it conditional probabilities for demographic factors  
30 20 such as age, income, family size and gender composition.  
The category 144 has associated with it conditional  
probabilities that represent probability that the viewing  
group is within a certain age group dependent on the  
35 probability that they are viewing a program in that category  
25 144.

40 FIG. 11 illustrates an entity-relationship diagram for  
the generation of the program demographic vectors 170. In a  
preferred embodiment, the heuristic rules 160 are applied  
along with the program characteristic vectors 150 in a  
30 30 program target analysis process 1100 to form the program  
demographic vectors 170. The program characteristic vectors  
150 indicate a particular aspect of a program, such as its  
violence level. The heuristic rules 160 indicate that a  
50 particular demographic group has a preference for that

5 program. As an example, it may be the case that young males  
have a higher preference for violent programs than other  
sectors of the population. Thus, a program which has the  
program characteristic vectors 150 indicating a high  
10 5 probability of having violent content, when combined with  
the heuristic rules 160 indicating that "young males like  
violent programs," will result, through the program target  
analysis process 1100, in the program demographic vectors  
15 170 which indicate that there is a high probability that the  
10 program is being watched by a young male.

The program target analysis process 1100 can be  
realized using software programmed in a variety of languages  
which processes mathematically the heuristic rules 160 to  
derive the program demographic vectors 170. The table  
15 representation of the heuristic rules 160 illustrated in  
FIG. 10B expresses the probability that the individual or  
25 household is from a specific demographic group based on a  
program with a particular category 144. This can be  
expressed, using probability terms as follow "the  
30 20 probability that the individuals are in a given demographic  
group conditional to the program being in a given category".  
Referring to FIG. 9D, the probability that the group has  
certain demographic characteristics based on the program  
35 being in a specific category is illustrated.

25 Expressing the probability that a program is destined  
to a specific demographic group can be determined by  
applying Bayes rule. This probability is the sum of the  
40 conditional probabilities that the demographic group likes  
the program, conditional to the category 144 weighted by the  
30 probability that the program is from that category 144. In a  
45 preferred embodiment, the program target analysis can  
calculate the program demographic vectors by application of  
logical heuristic rules, as illustrated in FIG. 10A, and by  
50 application of heuristic rules expressed as conditional

probabilities as shown in FIG. 10B. Logical heuristic rules can be applied using logical programming and fuzzy logic using techniques well understood by those skilled in the art, and are discussed in the text by S. V. Kartalopoulos entitled "Understanding Neural Networks and Fuzzy Logic" which is incorporated herein by reference.

Conditional probabilities can be applied by simple mathematical operations multiplying program context vectors by matrices of conditional probabilities. By performing this process over all the demographic groups, the program target analysis process 1100 can measure how likely a program is to be of interest to each demographic group. Those probabilities values form the program demographic vector 170 represented in FIG.12.

As an example, the heuristic rules expressed as conditional probabilities shown in FIG. 10B are used as part of a matrix multiplication in which the program characteristics vector 150 of dimension N, such as those shown in FIGS. 9A-9F is multiplied by an N x M matrix of heuristic rules expressed as conditional probabilities, such as that shown in FIG. 10B. The resulting vector of dimension M is a weighted average of the conditional probabilities for each category and represents the household demographic characteristics 190. Similar processing can be performed at the sub-category and content levels.

FIG. 12 illustrates an example of the program demographic vector 170, and shows the extent to which a particular program is destined to a particular audience. This is measured in terms of probability as depicted in FIG. 12. The Y-axis is the probability of appealing to the demographic group identified on the X-axis.

FIG. 13 illustrates an entity-relationship diagram for the generation of household session demographic data 1310 and household session interest profile 1320. In a preferred

embodiment, the subscriber selection data 110 is used along  
with the program characteristics vectors 150 in a session  
characterization process 1300 to generate the household  
session interest profile 1320. The subscriber selection data  
110 indicates what the subscriber is watching, for how long  
and at what volume they are watching the program.

In a preferred embodiment, the session characterization  
process 1300 forms a weighted average of the program  
characteristics vectors 150 in which the time duration the  
program is watched is normalized to the session time  
(typically defined as the time from which the unit was  
turned on to the present). The program characteristics  
vectors 150 are multiplied by the normalized time duration  
(which is less than one unless only one program has been  
viewed) and summed with the previous value. Time duration  
data, along with other subscriber viewing information, is  
available from the subscriber selection data 110. The  
resulting weighted average of program characteristics  
vectors forms the household session interest profile 1320,  
with each program contributing to the household session  
interest profile 1320 according to how long it was watched.  
The household session interest profile 1320 is normalized to  
produce probabilistic values of the household programming  
interests during that session.

In an alternate embodiment, the heuristic rules 160 are  
applied to both the subscriber selection data 110 and the  
program characteristics vectors 150 to generate the  
household session demographic data 1310 and the household  
session interest profile 1320. In this embodiment, weighted  
averages of the program characteristics vectors 150 are  
formed based on the subscriber selection data 110, and the  
heuristic rules 160 are applied. In the case of logical  
heuristic rules as shown in FIG. 10A, logical programming  
can be applied to make determinations regarding the

household session demographic data 1310 and the household session interest profile 1320. In the case of heuristic rules in the form of conditional probabilities such as those illustrated in FIG. 10B, a dot product of the time averaged values of the program characteristics vectors can be taken with the appropriate matrix of heuristic rules to generate both the household session demographic data 1310 and the household session interest profile 1320.

Volume control measurements which form part of the subscriber selection data 110 can also be applied in the session characterization process 1300 to form a household session interest profile 1320. This can be accomplished by using normalized volume measurements in a weighted average manner similar to how time duration is used. Thus, muting a show results in a zero value for volume, and the program characteristics vector 150 for this show will not be averaged into the household session interest profile 1320.

FIG. 14 illustrates an entity-relationship diagram for the generation of average household demographic characteristics and session household demographic characteristics 190. A household demographic characterization process 1400 generates the household demographic characteristics 190 represented in table format in FIG. 15. The household demographic characterization process 1400 uses the household viewing habits 195 in combination with the heuristic rules 160 to determine demographic data. For example, a household with a number of minutes watched of zero during the day may indicate a household with two working adults. Both logical heuristic rules as well as rules based on conditional probabilities can be applied to the household viewing habits 195 to obtain the household demographics characteristics 190.

The household viewing habits 195 is also used by the system to detect out-of-habits events. For example, if a

5 household with a zero value for the minutes watched column  
702 at late night presents a session value at that time via  
the household session demographic data 1310, this session  
will be characterized as an out-of-habits event and the  
10 5 system can exclude such data from the average if it is  
highly probable that the demographics for that session are  
greatly different than the average demographics for the  
household. Nevertheless, the results of the application of  
15 the household demographic characterization process 1400 to  
10 the household session demographic data 1310 can result in  
valuable session demographic data, even if such data is not  
added to the average demographic characterization of the  
20 household.

FIG. 15 illustrates the average and session household  
15 demographic characteristics. A household demographic  
parameters column 1501 is followed by an average value  
25 column 1505, a session value column 1503, and an update  
column 1507. The average value column 1505 and the session  
value column 1503 are derived from the household demographic  
30 20 characterization process 1400. The deterministic parameters  
such as address and telephone numbers can be obtained from  
an outside source or can be loaded into the system by the  
subscriber or a network operator at the time of  
35 installation. Updating of deterministic values is prevented  
25 by indicating that these values should not be updated in the  
update column 1507.

FIG. 16 illustrates an entity-relationship diagram for  
40 the generation of the household interest profile 180 in a  
household interest profile generation process 1600. In a  
30 preferred embodiment, the household interest profile  
45 generation process comprises averaging the household session  
interest profile 1320 over multiple sessions and applying  
the household viewing habits 195 in combination with the  
heuristic rules 160 to form the household interest profile  
50

180 which takes into account both the viewing preferences of the household as well as assumptions about households/subscribers with those viewing habits and program preferences.

FIG. 17 illustrates the household interest profile 180 which is composed of a programming types row 1709, a products types row 1707, and a household interests column 1701, an average value column 1703, and a session value column 1705.

The product types row 1707 gives an indication as to what type of advertisement the household would be interested in watching, thus indicating what types of products could potentially be advertised with a high probability of the advertisement being watched in its entirety. The programming types row 1709 suggests what kind of programming the household is likely to be interested in watching. The household interests column 1701 specifies the types of programming and products which are statistically characterized for that household.

As an example of the industrial applicability of the invention, a household will perform its normal viewing routine without being requested to answer specific questions regarding likes and dislikes. Children may watch television in the morning in the household, and may change channels during commercials, or not at all. The television may remain off during the working day, while the children are at school and day care, and be turned on again in the evening, at which time the parents may "surf" channels, mute the television during commercials, and ultimately watch one or two hours of broadcast programming. The present invention provides the ability to characterize the household, and may make the determination that there are children and adults in the household, with program and product interests indicated in the household interest profile 180 corresponding to a

5 family of that composition. A household with two retired  
adults will have a completely different characterization  
which will be indicated in the household interest profile  
180.

10 5 Although the present invention has been largely  
described in the context of a single computing platform  
receiving programming, the SCS 100 can be realized as part  
of a client-server architecture, as illustrated in FIG. 18.  
15 Referring to FIG. 18, residence 1800 contains a personal  
10 computer (PC) 1820 as well as the combination of a  
television 1810 and a set-top 1808, which can request and  
receive programming. The equipment in residence 1800, or  
20 similar equipment in a small or large business environment,  
forms the client side of the network as defined herein.  
15 Programming is delivered over an access network 1830, which  
may be a cable television network, telephone type network,  
25 or other access network. Information requests are made by  
the client side to a server 1840 which forms the server side  
of the network. Server 1840 has content locally which it  
30 provides to the subscriber, or requests content on behalf of  
the subscriber from a third party content provider 1860, as  
illustrated in FIG. 18. Requests made on behalf of the  
35 client side by server 1840 are made across a wide area  
network 1850 which can be the Internet or other public or  
25 private network. Techniques for making requests on behalf of  
a client are frequently referred to a proxy techniques and  
40 are well known to those skilled in the art. The server side  
receives the requested programming which is displayed on PC  
1820 or television 1810 according to which device made the  
30 request.

45 In a preferred embodiment the server 1840 maintains the  
subscriber selection data 110 which it is able to compile  
based on its operation as a proxy for the client side.  
50 Retrieval of source related information and the program



5 target analysis process 1100, the program characterization  
process 800, the program target analysis process 1100, the  
session characterization process 1300, the household  
demographic characterization process 1400, and the household  
10 5 interest profile generation process 1600 can be performed by  
server 1840.

Referring to FIG. 19 an advertisement monitoring table  
is illustrated, in which an advertisement ID (AD ID) column  
15 1915 contains a numerical ID for an advertisement which was  
10 transmitted with the advertisement in the form of a Program  
ID, http address, or other identifier which is uniquely  
associated with the advertisement. A product column 1921  
20 contains a product description which indicates the type of  
product that was advertised. A brand column 1927 indicates  
15 the brand name of the product or can alternatively list a  
generic name for that product. A percent watched column 1933  
25 indicates the percentage of the advertisement the subscriber  
viewed. In an alternate embodiment, a letter rating or  
other type of rating is used to indicate the probability  
30 20 that the advertisement was watched. A volume column 1937  
indicates the volume level at which the advertisement was  
watched.

As an example of the industrial applicability of the  
35 invention, a manufacturer may develop an advertising  
25 strategy which includes the insertion of advertisements  
during popular evening programs. The costs for such ad  
40 insertions can be extremely high. In order to insure the  
cost effectiveness of this advertising strategy, the  
manufacturer has the advertisements placed during less  
30 30 watched but similar programs and monitors how subscribers  
45 react, and can determine approximately how many times the  
advertisement has been watched out of all of the possible  
viewings. This data can be used to confirm the potential  
50 effectiveness of the advertisement and to subsequently

5 determine if purchasing the more expensive time during  
evening programming will be cost-effective, or if the  
advertisement should be modified or placed in other  
programming.

10 5 Continuing this example, the manufacturer may place an  
advertisement for viewing during "prime time" for an initial  
period but can subsequently cancel broadcasts of the  
advertisement if it is found that the majority of  
15 subscribers never see the advertisement.

10 10 Although this invention has been illustrated by  
reference to specific embodiments, it will be apparent to  
those skilled in the art that various changes and  
20 modifications may be made which clearly fall within the  
scope of the invention. The invention is intended to be  
15 protected broadly within the spirit and scope of the  
appended claims.

## Claims

5

10

15

20

25

30

35

40

45

50

55

## Claims

5

What is claimed is:

10

1. A data processing system for monitoring advertisements watched by a subscriber, said data processing system comprising:

15

(a) computer processor means for processing data;  
(b) storage means for storing data on a storage medium;

20

(c) first means for monitoring subscriber activity wherein said first means includes recording means for storing subscriber selections;

25

(d) second means for retrieving advertisement related information wherein said advertisement related information contains descriptive fields corresponding to said advertisement;

30

(e) third means for processing information wherein said third means includes means for determining the extent to which an advertisement is viewed by said subscriber; and

35

(f) fourth means for storing said descriptive fields and said determination of the extent to which said advertisement is viewed by said subscriber.

40

2. The system described in claim 1 wherein said first means for monitoring subscriber activity further comprises means for monitoring volume levels wherein said volume levels correspond to subscriber selection volume levels.

45

3. The system described in claim 1 further comprising:

50

(g) fifth means for determining a subscriber product interests profile; and

55

5

(h) sixth means for storing said subscriber products interests profile.

10

4. The system described in claim 1 wherein said second means for retrieving advertisement related information further comprises a means for context mining of textual information associated with said selected source material.

15

5. The system described in claim 4 wherein said textual information is text derived from closed-captioning data associated with said advertisement.

20

6. The system described in claim 5 wherein said text derived from closed-captioning data associated with said advertisement includes a product name field.

25

7. The system described in claim 4 wherein said text derived from closed-captioning data associated with said advertisement includes a product brand field.

30

8. A client-server based data processing system for monitoring advertisements watched by a subscriber, said client-server based data processing system comprising:

35

(a) first computer processor means at a client side for receiving and displaying advertisements wherein said first computer means is capable of transmitting channel change requests;

40

(b) second computer processor means at a server side for receiving said channel change requests and for processing data;

45

(c) second storage means associated with second computer processor means for storing data on a storage medium;

50

55

5 (d) first means at said server side for  
monitoring subscriber activity wherein said  
first means for monitoring subscriber activity  
includes receiving means for receiving  
10 subscriber channel change requests, recording  
means for storing subscriber channel change  
requests;

15 (e) second means at said server side for  
retrieving advertisement related information  
wherein said advertisement related information  
contains descriptive fields corresponding to an  
advertisement;

20 (f) third means at said server side for  
processing information wherein said third means  
includes means for determining the extent to  
25 which an advertisement is viewed by said  
subscriber; and

(g) fourth means at said server side for storing  
said descriptive fields and said determination  
30 of the extent to which said advertisement is  
viewed by said subscriber.

35 9. The system described in claim 8 further comprising:

(h) fifth means for determining a subscriber  
product interests profile; and  
40 (i) sixth means for storing said subscriber  
product interests profile.

45 10. The system described in claim 8 wherein said second  
means for retrieving advertisement related information  
further comprises a means for context mining of textual  
information associated with said selected source material.

50

55

5 11. The system described in claim 10 wherein said textual information is text derived from closed-captioning data associated with said advertisement.

10 12. The system described in claim 11 wherein said text derived from closed-captioning data associated with said advertisement includes a product name field.

15 13. The system described in claim 11 wherein said text derived from closed-captioning data associated with said advertisement includes a product brand field.

20 14. A data processing system for generating a subscriber profile vector, said data processing system comprising:

- 25 (a) computer processor means for processing data;
- (b) storage means for storing data on a storage medium;
- 30 (c) first means for monitoring subscriber activity wherein said first means includes recording means for storing subscriber selection data wherein said subscriber selection data corresponds to selected source material;
- 35 (d) second means for retrieving source related information wherein said source related information contains descriptive fields corresponding to said selected source material;
- 40 (e) third means for processing information wherein said third means includes means for processing said subscriber selection data with respect to said descriptive fields to form said
- 45 subscriber profile vector; and
- (f) fourth means for storing said subscriber profile vector.

5           15. The system described in claim 14 wherein said first  
means for monitoring subscriber activity further comprises  
means for monitoring time durations wherein said time  
durations correspond to viewing times of said selected  
10 source material.

16. The system described in claim 14 wherein said first  
means for monitoring subscriber activity further comprises  
15 means for monitoring volume levels wherein said volume  
levels correspond to subscriber selection volume levels.

20           17. The system described in claim 14 wherein said  
subscriber profile vector contains household demographic  
data indicating probabilistic measurements of household  
demographics.

25           18. The system described in claim 14 wherein said  
subscriber profile vector contains household program  
preference information indicating probabilistic measurements  
30 of household program interests.

35           19. The system described in claim 14 wherein said  
subscriber profile vector contains household product  
preference information indicating probabilistic measurements  
of household product interests.

40           20. The system described in claim 14 wherein said second  
means for retrieving source related information further  
comprises a means for context mining of textual information  
associated with said selected source material.

45           21. The system described in claim 20 wherein said textual  
information is text derived from closed-captioning data  
associated with said selected source material.  
50



5           22. The system described in claim 14 wherein said second  
means for retrieving source related information further  
comprises a means for retrieving information associated with  
10       said selected source material from an electronic program  
guide.

15           23. The system described in claim 14 wherein said third  
means for processing information processes information over  
a viewing session and wherein said subscriber profile vector  
corresponds to said viewing session.

20           24. The system described in claim 14 wherein said third  
means for processing information processes information over  
multiple viewing sessions and wherein said subscriber  
25       profile vector corresponds to an average value over said  
multiple viewing sessions.

30           25. A data processing system for generating a subscriber  
profile vector, said data processing system comprising:

- (a) computer processor means for processing data;
- (b) storage means for storing data on a storage  
35       medium;
- (c) first means for monitoring subscriber  
activity wherein said first means includes  
recording means for storing subscriber selection  
40       data wherein said subscriber selection data  
corresponds to selected source material;
- (d) second means for retrieving source related  
information wherein said source related  
45       information contains descriptive fields  
corresponding to said selected source material;

50

55

5 (e) third means for generating a program characteristics vector based on said source related information;

(f) fourth means for storing a set of heuristic rules;

10 (g) fifth means for processing information wherein said fifth means includes means for processing said subscriber selection data with respect to said program characteristics vector and said set of heuristic rules to form said subscriber profile vector; and

15 (h) sixth means for storing said subscriber profile vector.

20 26. The system described in claim 25 wherein said first means for monitoring subscriber activity further comprises means for monitoring time durations wherein said time durations correspond to viewing times of said selected source material.

25 27. The system described in claim 25 wherein said first means for monitoring subscriber activity further comprises means for monitoring volume levels wherein said volume levels correspond to subscriber selection volume levels.

30 28. The system described in claim 25 wherein said subscriber profile vector contains household demographic data indicating probabilistic measurements of household demographics.

35 29. The system described in claim 25 wherein said subscriber profile vector contains a household session interest profile indicating probabilistic measurements of household interests.

5           30. A data processing system for generating a household  
demographic characteristics vector, said data processing  
system comprising:

- 10           (a) computer processor means for processing data;  
            (b) storage means for storing data on a storage  
                medium;  
            (c) first means for monitoring subscriber  
15           activity wherein said first means includes  
            recording means for storing subscriber selection  
            data wherein said subscriber selection data  
            corresponds to selected source material;  
20           (d) second means for generating household viewing  
            habits information wherein said household  
            viewing habits information is generated from  
25           said subscriber selection data;  
            (e) third means for storing a set of heuristic  
                rules;  
            (f) fourth means for processing information  
30           wherein said fourth means includes means for  
            processing said subscriber selection data with  
            respect to said set of heuristic rules to form  
            said household demographic characteristics  
35           vector; and  
            (g) fifth means for storing said household  
                demographic characteristics vector.

40           31. The system described in claim 30 wherein said fourth  
means for processing information processes information over  
a viewing session and wherein said household demographic  
45           characteristics vector corresponds to said viewing session.

50           32. The system described in claim 30 wherein said fourth  
means for processing information processes information over

5 a period of multiple viewing sessions wherein said household demographic characteristics vector corresponds to an average value over said multiple viewing sessions.

10 33. A data processing system for generating a subscriber profile vector in a client-server based architecture, said data processing system comprising:

15 (a) first computer processor means at a client side for requesting and displaying source information wherein said first computer means transmits a request for source material and receives and displays said source material;

20 (b) second computer processor means at a server side for processing data;

25 (c) second storage means associated with second computer processor means for storing data on a storage medium;

30 (d) first means at said server side for monitoring subscriber activity wherein said first means for monitoring subscriber activity includes receiving means for receiving subscriber requests for said source material, recording means for storing subscriber selection data wherein said subscriber selection data corresponds to a record of requests for said source material;

35 (e) second means at said server side for retrieving source related information wherein said source related information contains descriptive fields corresponding to said source material;

40 (f) third means at said server side for processing information wherein said third means includes means for processing said subscriber

5                   selection data with respect to said descriptive  
                  fields to form said subscriber profile vector;  
                  and

10                   (g) fourth means at said server side for storing  
                  said subscriber profile vector.

15                   34. The system described in claim 33 wherein said first  
                  means for monitoring subscriber activity further comprises  
                  means for monitoring time durations wherein said time  
                  durations correspond to viewing times of said selected  
                  source material.

20                   35. The system described in claim 33 wherein said first  
                  means for monitoring subscriber activity further comprises  
                  means for monitoring volume levels wherein said volume  
25                   levels correspond to subscriber selection volume levels.

30                   36. The system described in claim 33 wherein said  
                  subscriber profile vector contains household demographic  
                  data indicating probabilistic measurements of household  
                  demographics.

35                   37. The system described in claim 33 wherein said  
                  subscriber profile vector contains household program  
                  preference information indicating probabilistic measurements  
                  of household program interests.

40                   38. The system described in claim 33 wherein said  
                  subscriber profile vector contains household product  
                  preference information indicating probabilistic measurements  
45                   of household product interests.

50                   39. The system described in claim 33 wherein said second  
                  means for retrieving source related information further

comprises a means for context mining of textual information associated with said selected source material.

40. The system described in claim 39 wherein said textual information is text derived from closed-captioning data associated with said selected source material.

41. The system described in claim 33 wherein said second means for retrieving source related information further comprises a means for retrieving information associated with said selected source material from an electronic program guide.

42. The system described in claim 33 wherein said third means for processing information processes information over a viewing session and wherein said subscriber profile vector corresponds to said viewing session.

43. The system described in claim 33 wherein said third means for processing information processes information over multiple viewing sessions and wherein said subscriber profile vector corresponds to an average value over said multiple viewing sessions.

44. A data processing system for generating a subscriber profile vector in a client-server based architecture, said data processing system comprising:

- (a) first computer processor means at a client side for requesting and displaying source information wherein said first computer means transmits a request for source material and receives and displays said source material;
- (b) second computer processor means at a server side for processing data;

5 (c) second storage means associated with second  
computer processor means for storing data on a  
storage medium;

10 (d) first means at said server side for  
monitoring subscriber activity wherein said  
first means for monitoring subscriber activity  
includes receiving means for receiving  
15 subscriber requests for said source material,  
recording means for storing subscriber selection  
data wherein said subscriber selection data  
corresponds to a record of requests for said  
source material;

20 (e) second means at said server side for  
retrieving source related information wherein  
said source related information contains  
25 descriptive fields corresponding to said source  
material;

(f) third means at said server side for  
generating a program characteristics vector  
30 based on said source related information;

(g) fourth means at said server side for storing  
a set of heuristic rules;

35 (h) fifth means at said server side for  
processing information wherein said fifth means  
includes means for processing said subscriber  
selection data with respect to said program  
40 characteristics vector and said set of heuristic  
rules to form said subscriber profile vector;  
and

45 (i) sixth means at said server side for storing  
said subscriber profile vector.

45. The system described in claim 44 wherein said first  
50 means for monitoring subscriber activity further comprises

5

means for monitoring time durations wherein said time durations correspond to viewing times of said selected source material.

10

46. The system described in claim 44 wherein said first means for monitoring subscriber activity further comprises means for monitoring volume levels wherein said volume levels correspond to subscriber selection volume levels.

15

20

47. The system described in claim 44 wherein said subscriber profile vector contains household demographic data indicating probabilistic measurements of household demographics.

25

48. The system described in claim 44 wherein said subscriber profile vector contains a household session interest profile indicating probabilistic measurements of household interests.

30

49. A data processing system for generating a household demographic characteristics vector in a client-server based architecture, said data processing system comprising:

35

(a) first computer processor means at a client side for requesting and displaying source information wherein said first computer means transmits a request for source material and receives and displays said source material;

40

(b) second computer processor means at a server side for processing data;

45

(c) first means at said server side for monitoring subscriber activity wherein said first means includes recording means for storing subscriber selection data wherein said

50

55



5 subscriber selection data corresponds to  
selected source material;

(d) second means at said client side for  
generating household viewing habits information  
10 wherein said household viewing habits  
information is generated from said subscriber  
selection data;

(e) third means at said server side for storing a  
15 set of heuristic rules;

(f) fourth means at said server side for  
processing information wherein said fourth means  
includes means for processing said subscriber  
20 selection data with respect to said set of  
heuristic rules to form said household  
demographic characteristics vector; and

(g) fifth means at said server side for storing  
25 said household demographic characteristics  
vector.

30 50. The system described in claim 49 wherein said fourth  
means for processing information processes information over  
a viewing session and wherein said household demographic  
characteristics vector corresponds to said viewing session.  
35

51. The system described in claim 49 wherein said fourth  
means for processing information processes information over  
40 a period of multiple viewing sessions wherein said household  
demographic characteristics vector corresponds to an average  
value over said multiple viewing sessions.

45

50

55

1/22

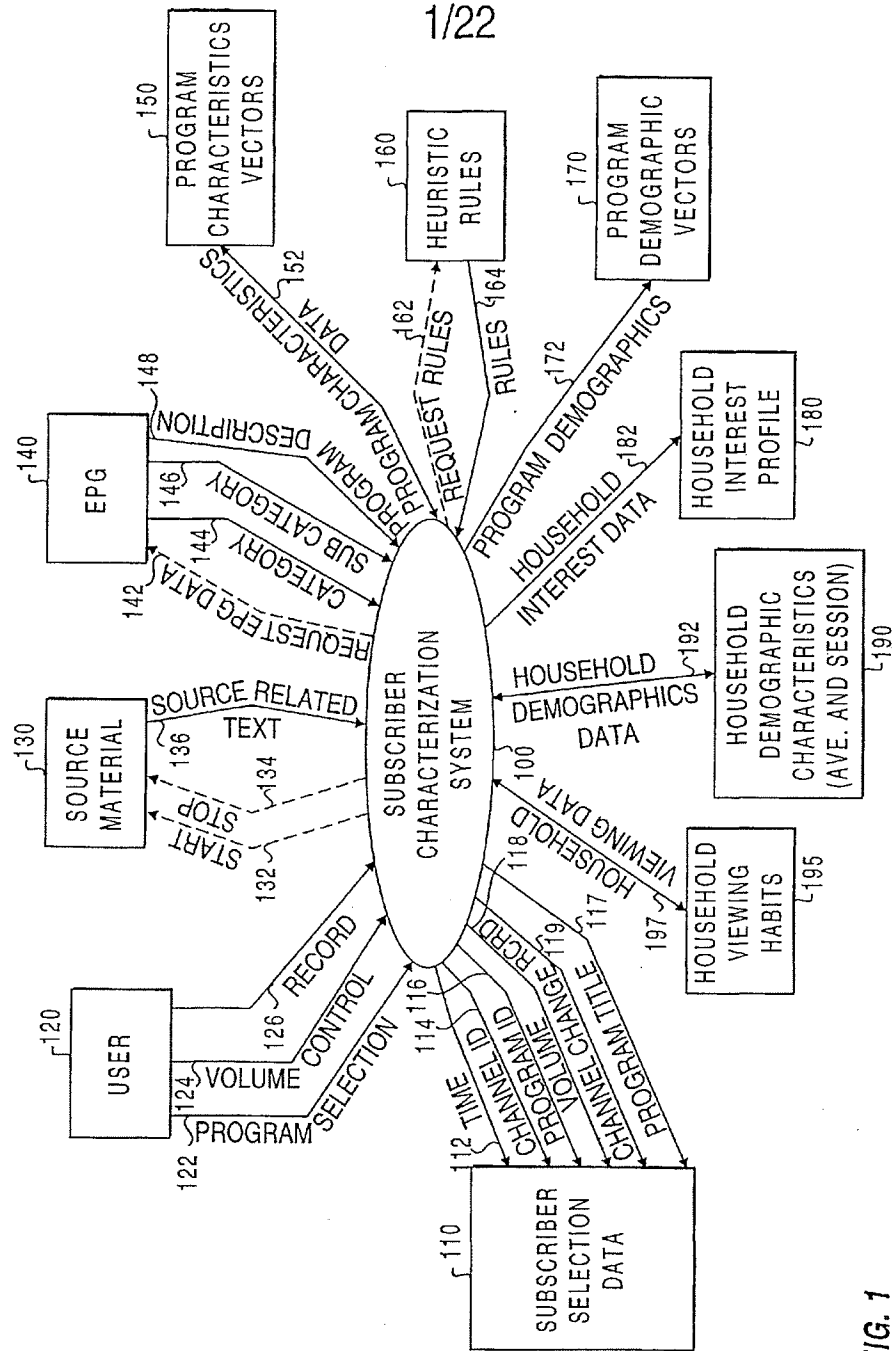


FIG. 1

2/22

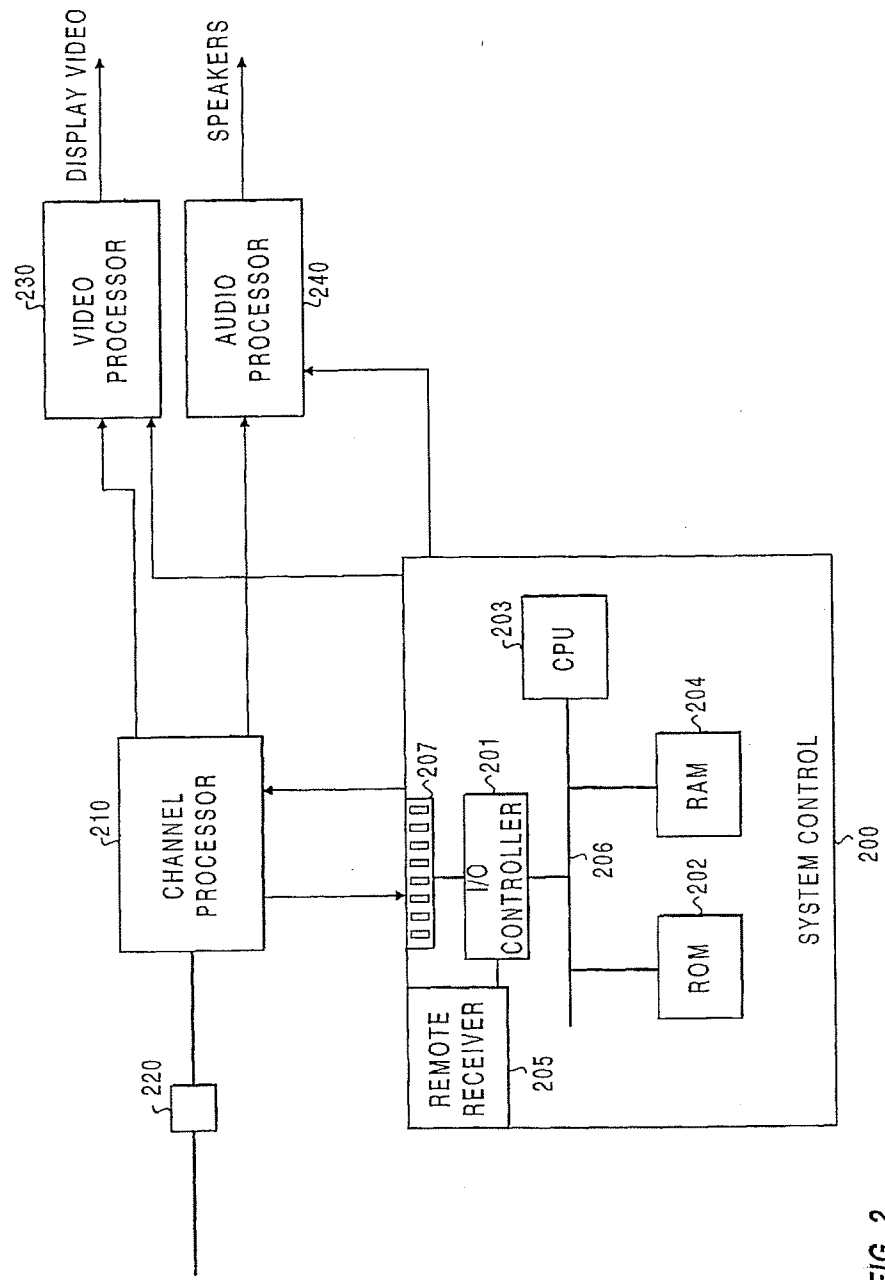


FIG. 2

3/22

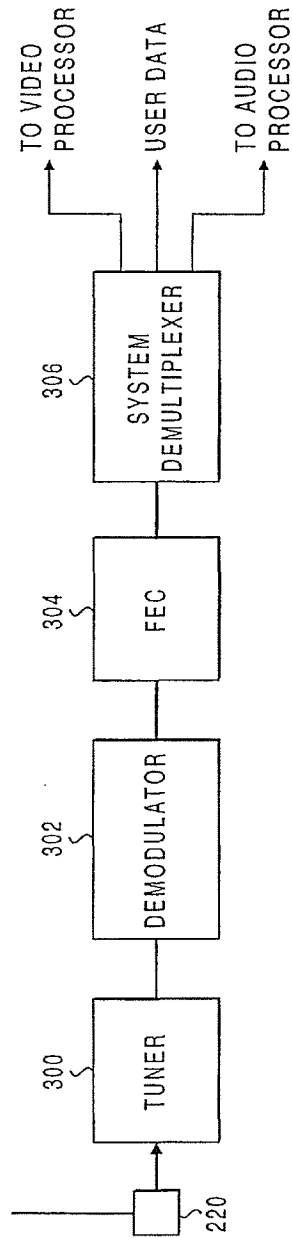


FIG. 3

4/22

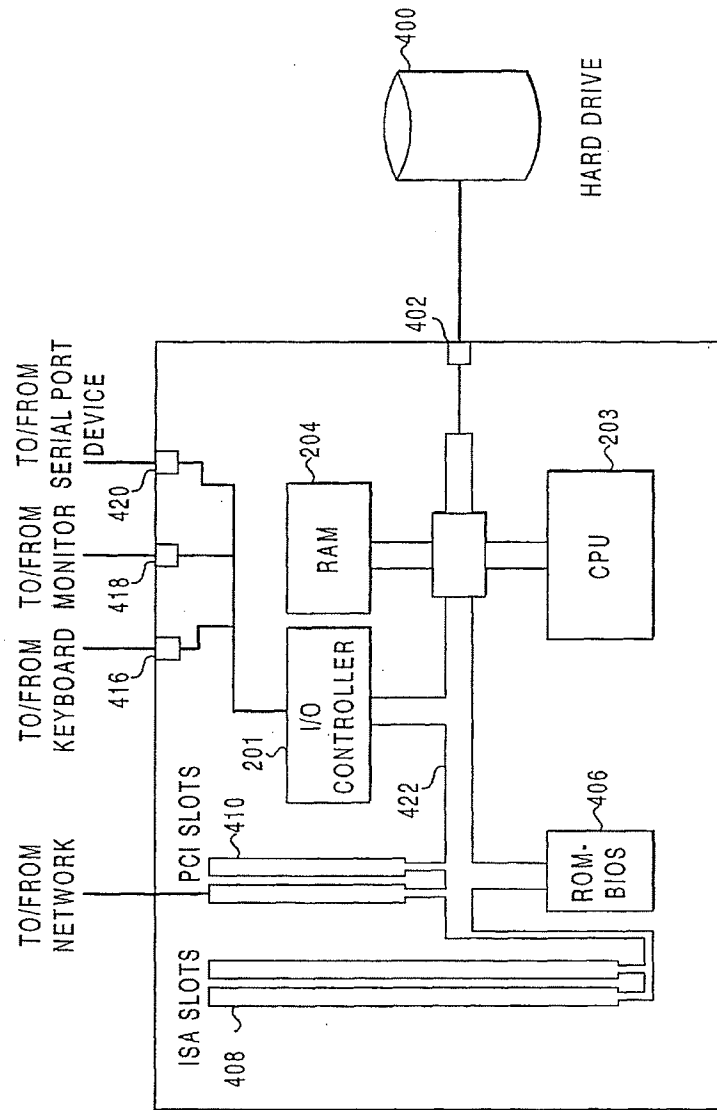


FIG. 4

5/22

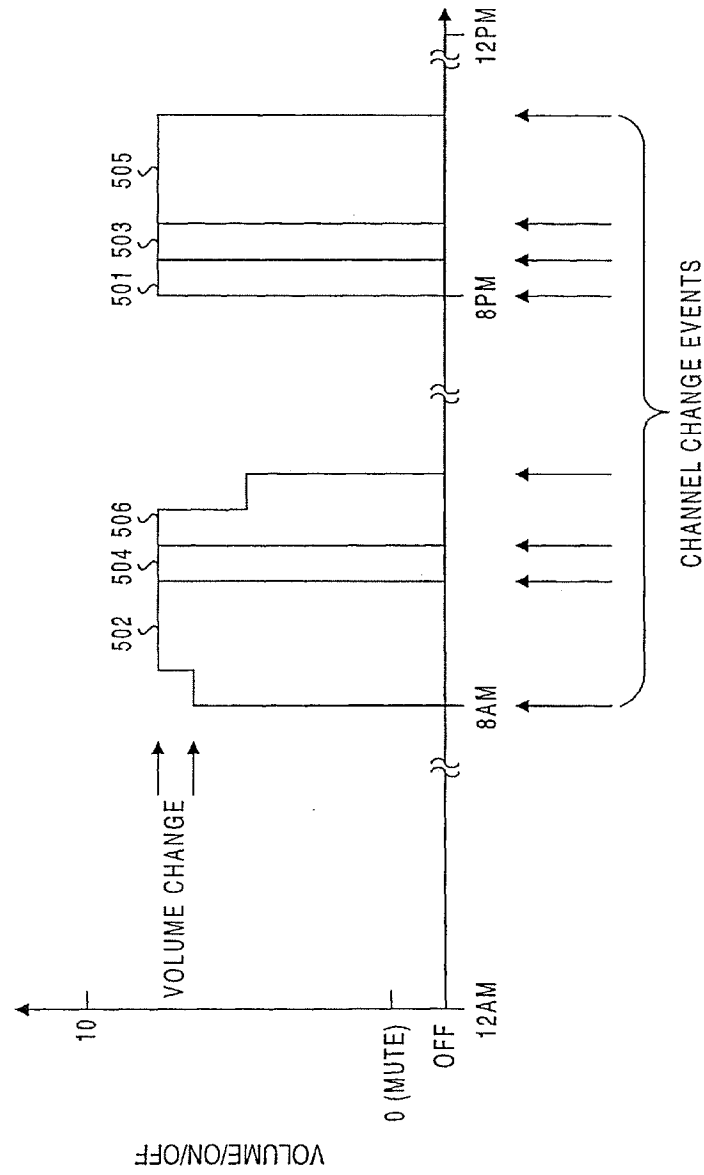


FIG. 5

6/22

602	604	603	601
TIME	CHANNEL ID	PROGRAM TITLE	VOLUME
08:01:25AM	06	"MORNING TV" "GOOD MORNING AMERICA" "GOOD MORNING AMERICA"	5/10
08:01:45AM	13		5/10
08:03:25AM	13		6/10
⋮			
06:11:25PM	09	"SEINFELD" "ADVERTISING" "SEINFELD" "ADVERTISING" "LIVING SINGLE"	5/10
06:15:23PM	09		5/10
06:17:25PM	09		5/10
06:28:10PM	09		5/10
06:30:07PM	52		5/10
⋮			

FIG. 6

7/22

TIME OF DAY	MINUTES WATCHED	CHANNEL CHANGES	AVERAGE VOLUME
MORNING (6AM-9AM)	61	2	5/10
MID-DAY (9AM-3PM)	0	0	-
AFTERNOON (3PM-6PM)	0	0	-
NIGHT (6PM-10PM)	122	4	6/10
LATE NIGHT (12AM-6AM)	0	0	-
TOTAL	183	6	5.7/10

FIG. 7



8/22

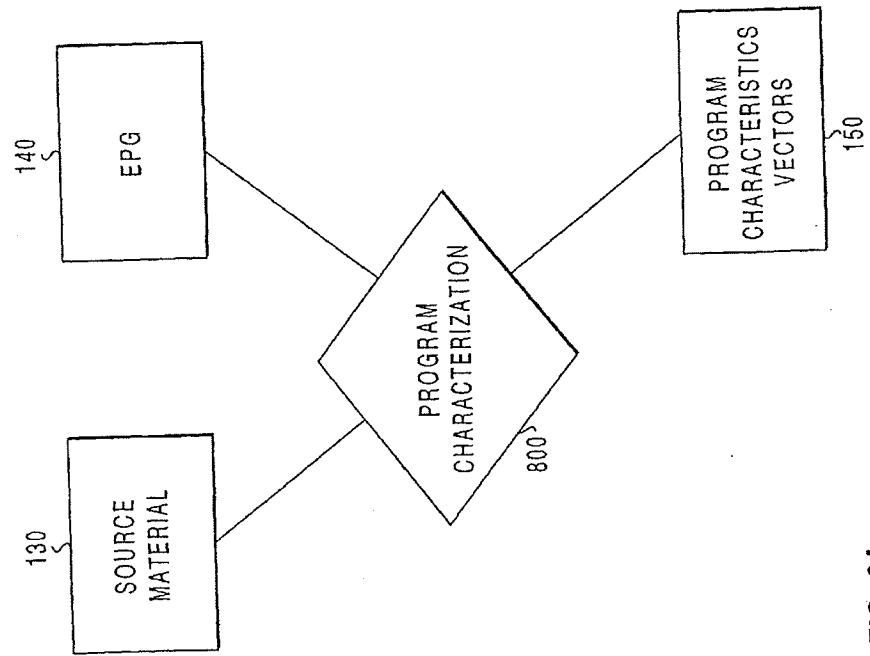


FIG. 8A

9/22

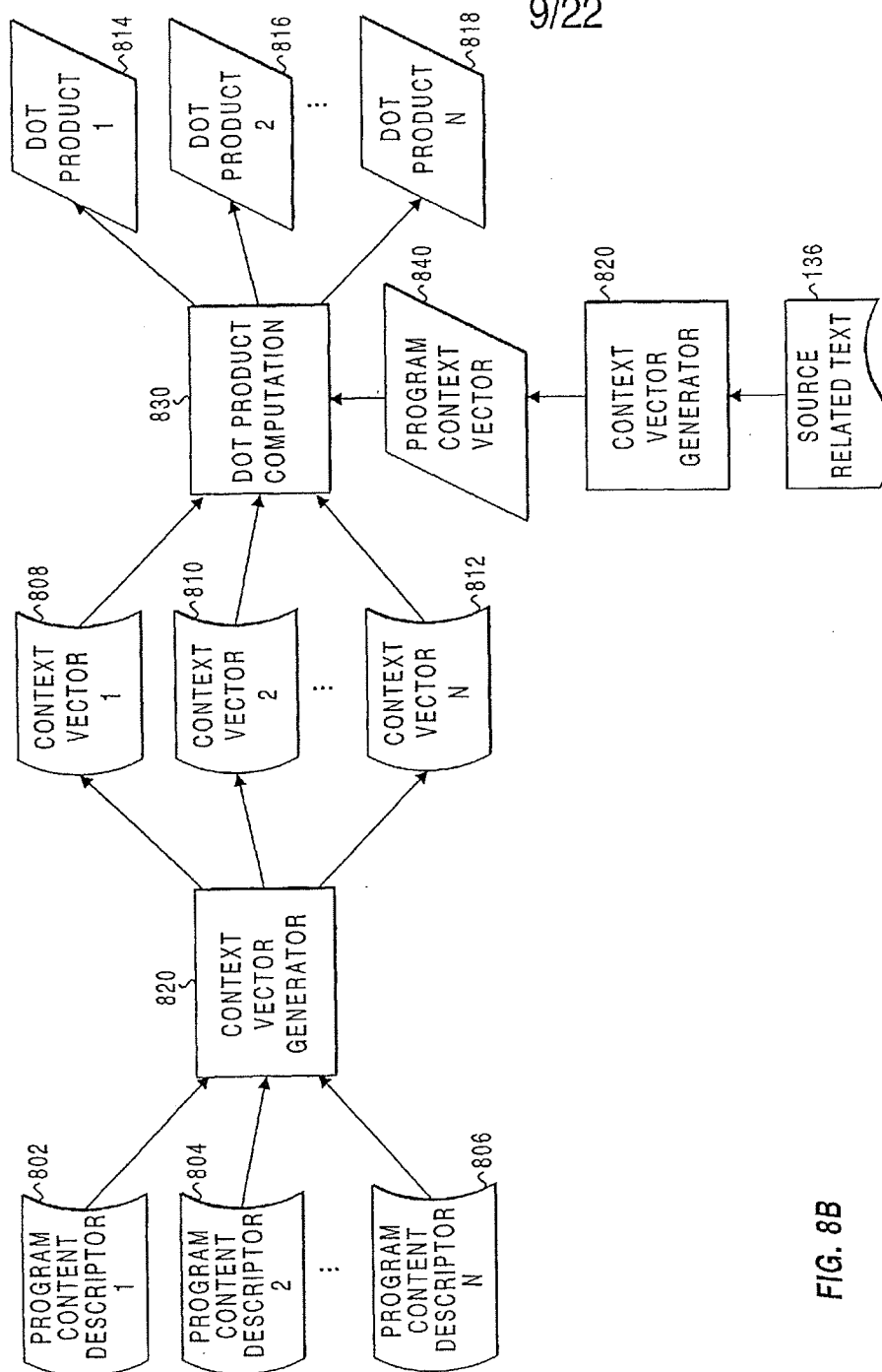


FIG. 8B

10/22

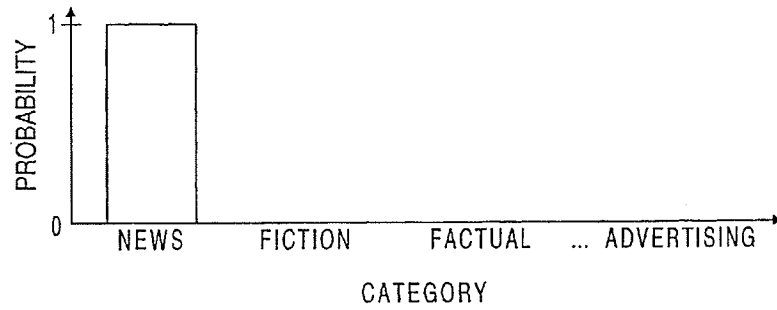


FIG. 9A

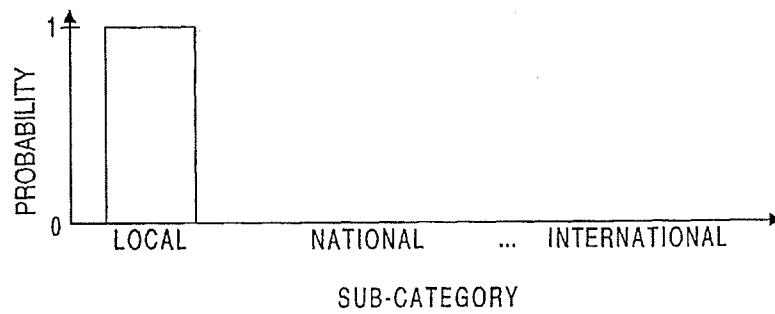


FIG. 9B

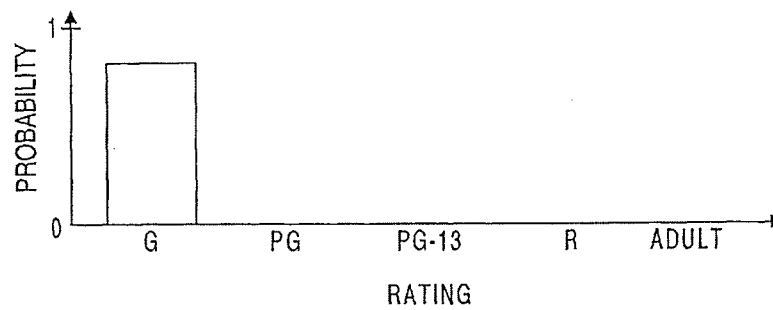


FIG. 9C

11/22

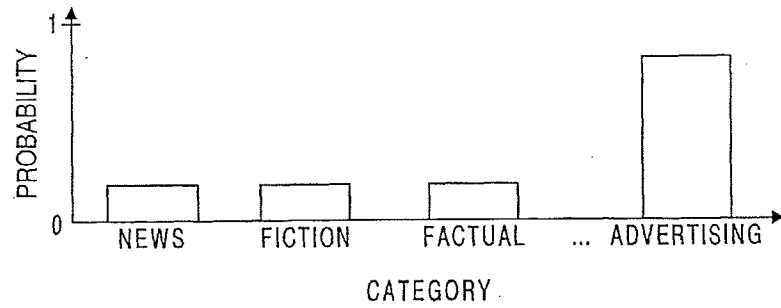


FIG. 9D

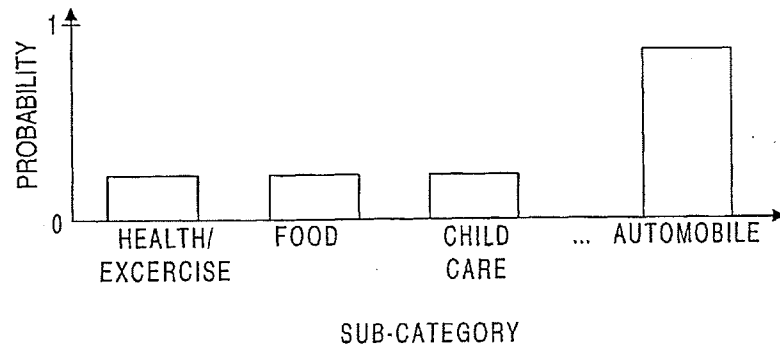


FIG. 9E

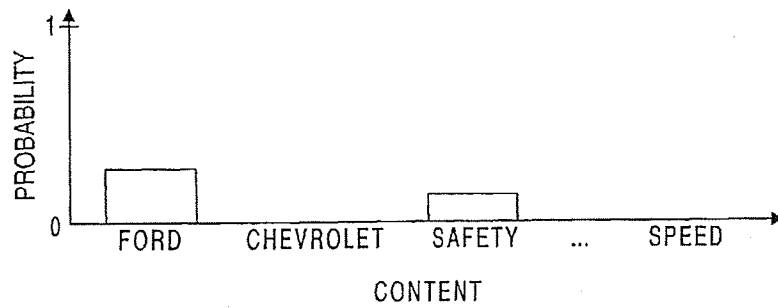


FIG. 9F

12/22

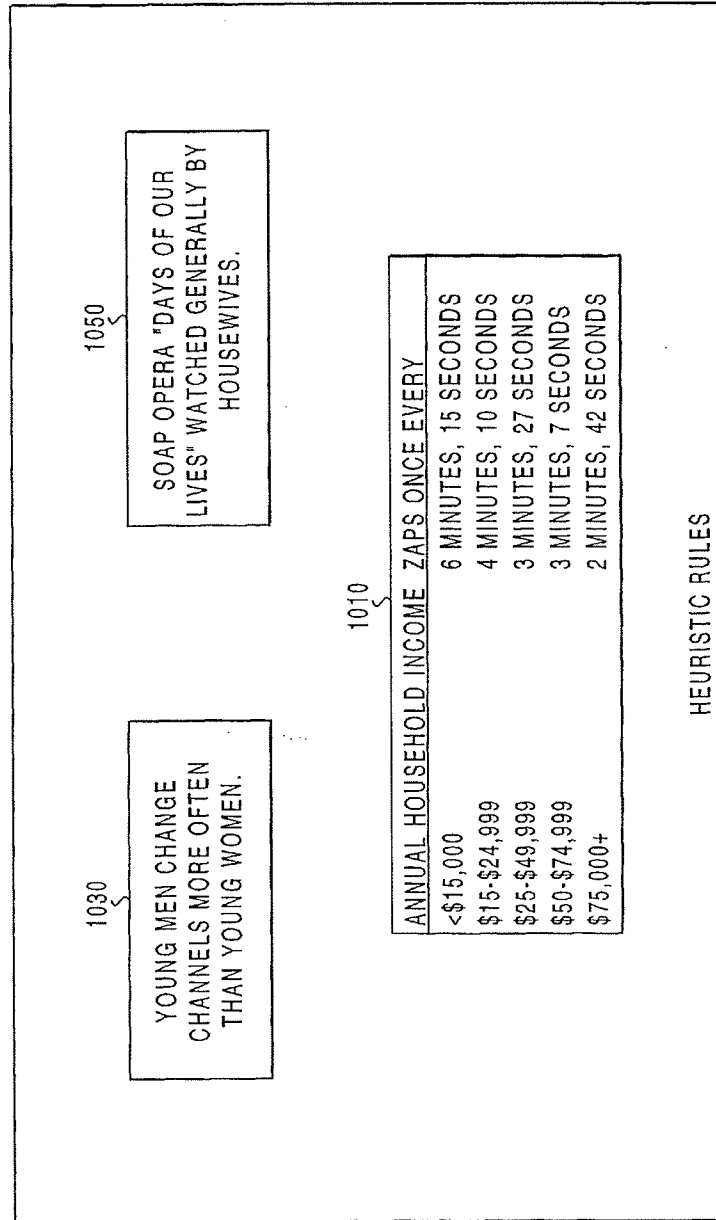


FIG. 10A

13/22

DEMOGRAPHIC GROUPS												
		AGE			INCOME			SIZE			GENDER	
		0-10	10-18	... >70	0-20K	20-50K	... 50-100K	1	2	... >5	M	F
CATEGORIES	NEWS	0.1	0.1	0.4	0.2	0.3	0.4	0.5	0.3	0.1	0.3	0.7
	FICTION	0.5	0.3	0.2	0.4	0.2	0.3	0.3	0.2	0.1	0.8	0.2
	FACTUAL	0.2	0.2	0.3	0.1	0.4	0.2	0.2	0.2	0.4	0.4	0.6
	⋮											
	ADVERTISING	0.1	0.3	0.5	0.3	0.2	0.1	0.2	0.1	0.3	0.5	0.5

FIG. 10B

14/22

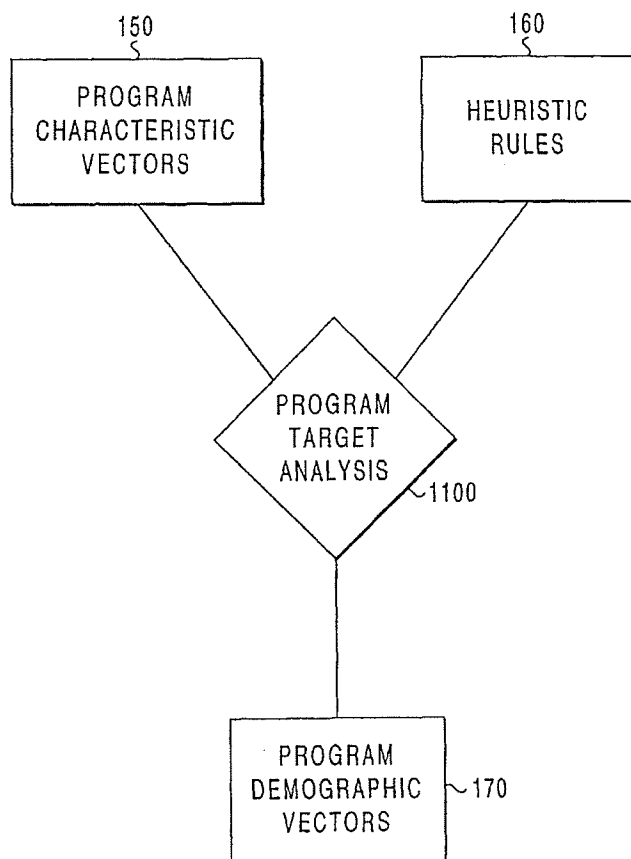


FIG. 11

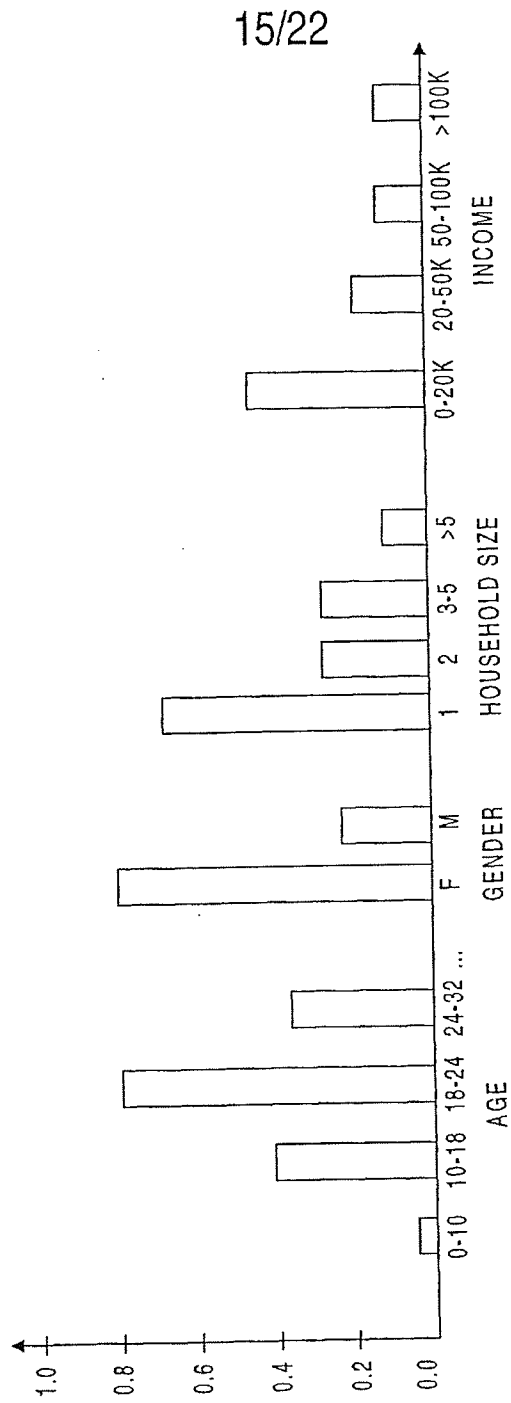


FIG. 12



16/22

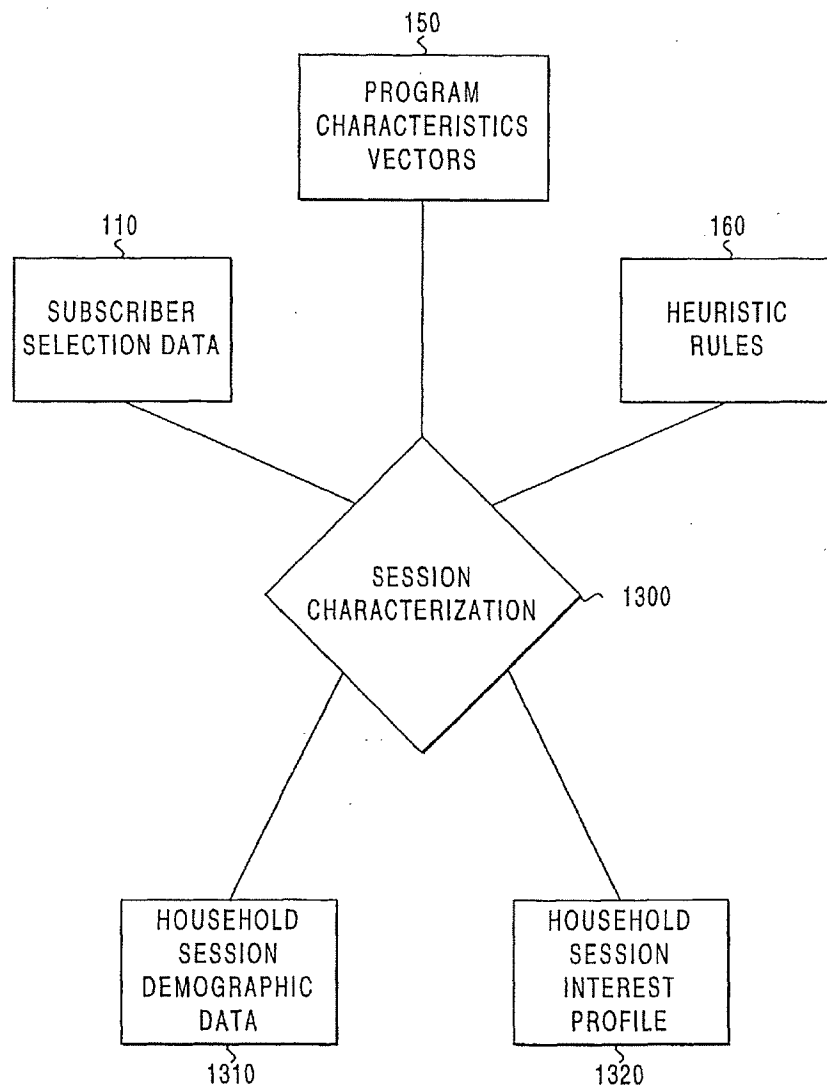


FIG. 13

17/22

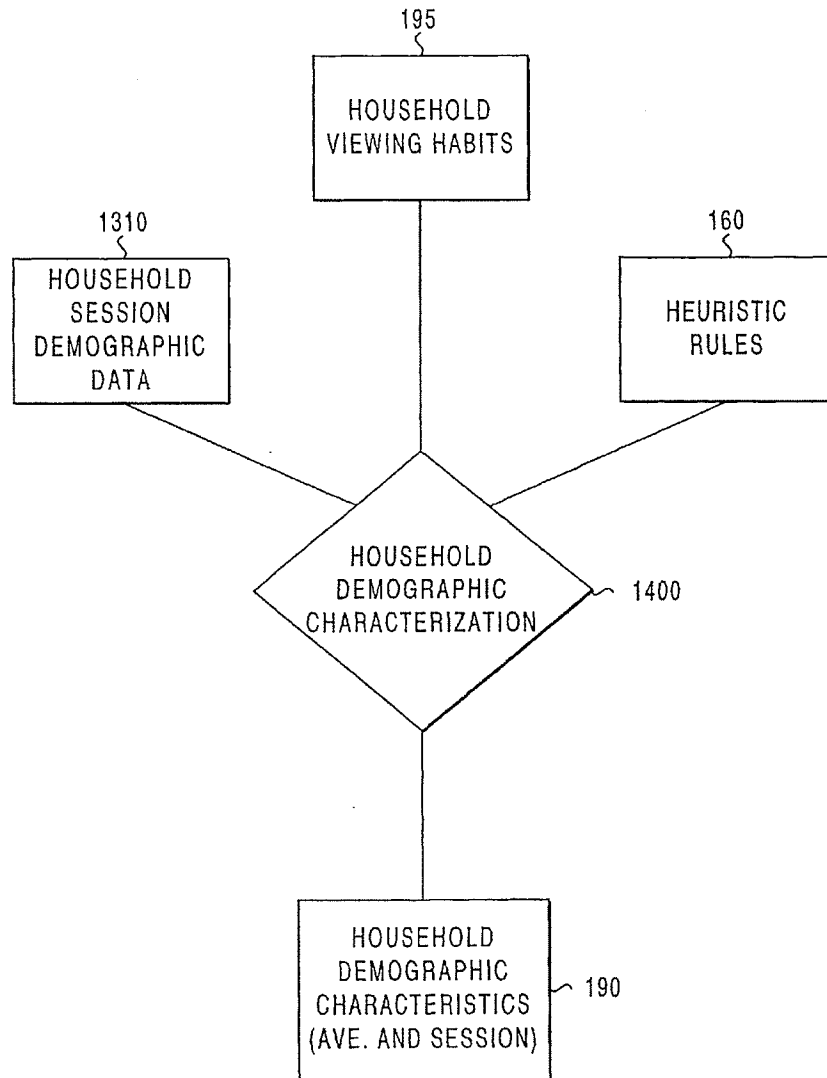


FIG. 14

18/22

1501 S	1505 S	1503 S	1507 S
HOUSEHOLD PARAMETER	AVERAGE VALUE	SESSION VALUE	UPDATE?
SIZE	2.6	3.0	YES
AGE	23.5	12	YES
SEX (FEMALE=1)	0.6	0.7	YES
INCOME (\$0-\$20K)	0.1	0.1	YES
INCOME (\$20-\$50K)	0.6	0.7	YES
INCOME (\$50-\$100K)	0.2	0.1	YES
INCOME (>\$100K)	0.1	0.1	YES
ZIP CODE			NO
TELEPHONE NUMBER			NO

FIG. 15

19/22

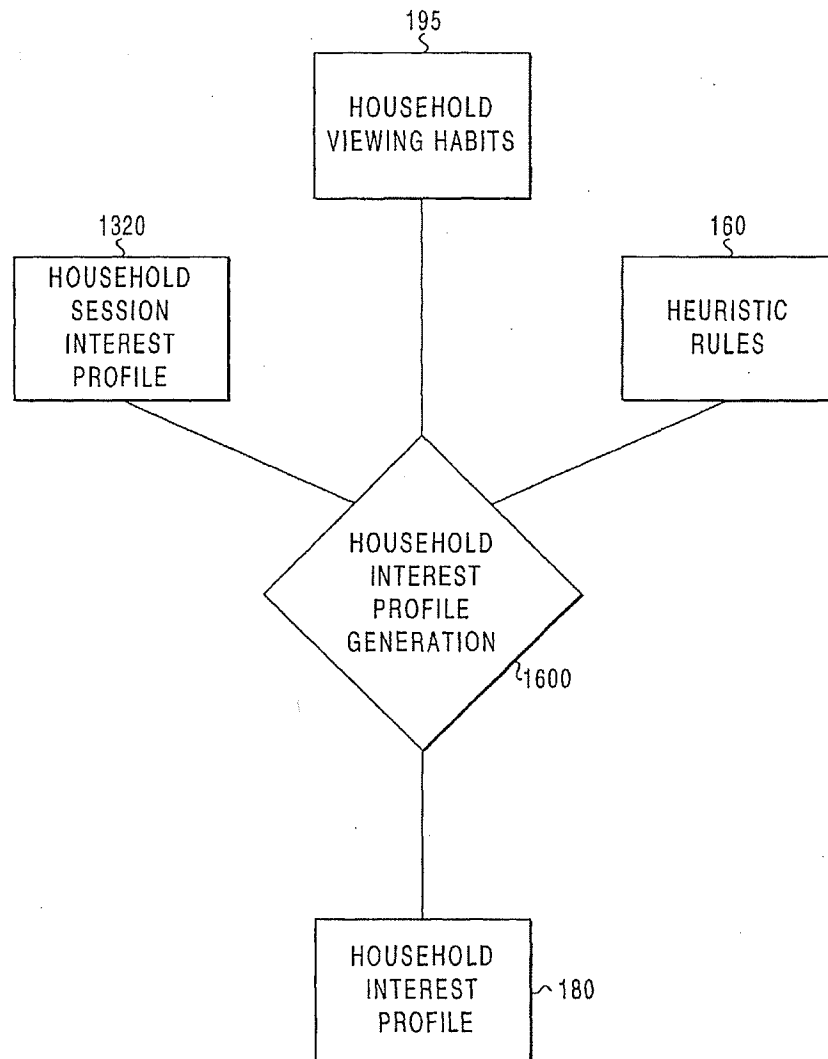


FIG. 16

20/22

		1701 \$	1703 \$	1705 \$
		HOUSEHOLD INTEREST	AVERAGE VALUE	SESSION VALUE
1709	PROGRAMMING	DRAMA	0.1	0.20
		ROMANCE	0.1	0.20
		ACTION	0.6	0.25
		SITCOM	0.2	0.30
		⋮		
		SPORTS	0	0.05
1707	PRODUCTS	HEALTH/EXCERCISE	0.6	0.2
		FOOD	0.3	0.4
		CHILD RELATED	0.0	0.1
		TOYS	0.0	0.1
		⋮		
		AUTOMOBILE	0.1	0.2

FIG. 17

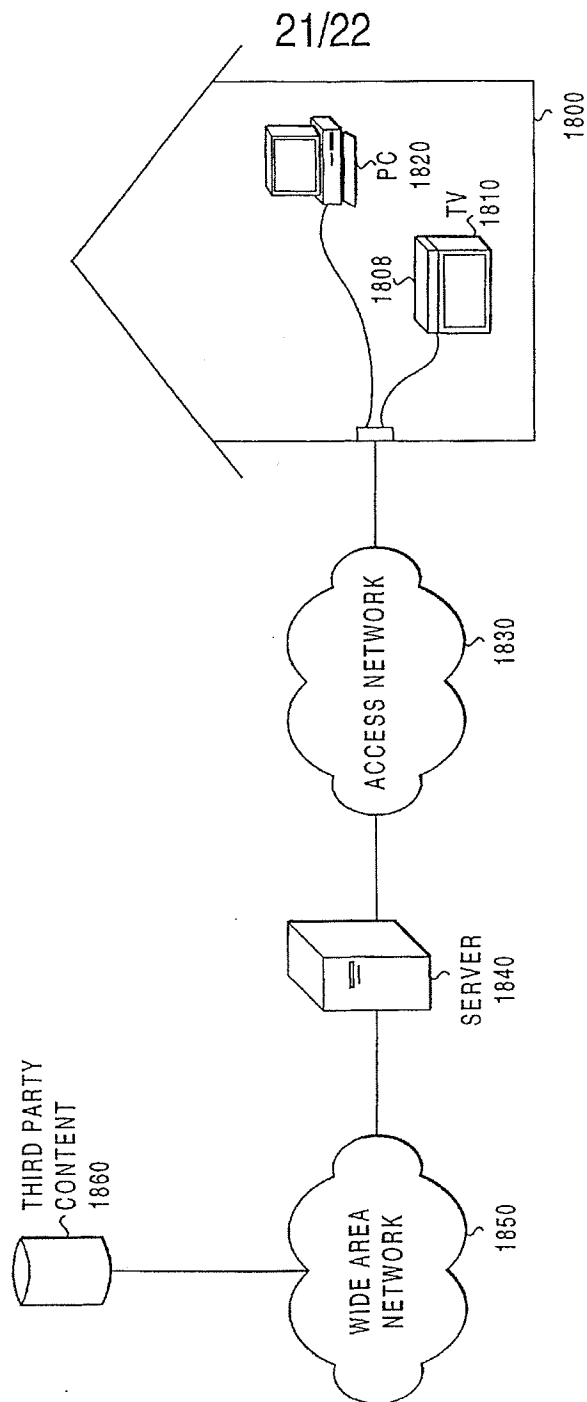


FIG. 18

22/22

<sup>1915</sup> \$ ADID	<sup>1921</sup> \$ PRODUCT	<sup>1927</sup> \$ BRAND	<sup>1933</sup> \$ % WATCHED	<sup>1937</sup> \$ VOLUME
216	DIAPERS	HUGS	50%	6/10
1230	DETERGENT	SOAPY	90%	6/10
	:			
4137	AUTOMOBILES	SPEEDSTER	70%	8/10

FIG. 19

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



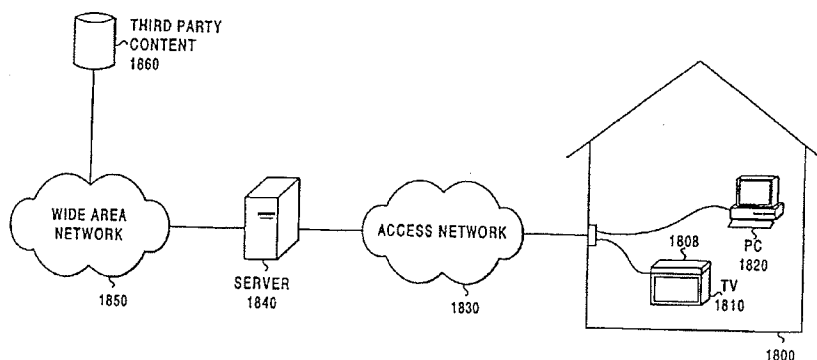
(43) International Publication Date  
8 June 2000 (08.06.2000)

PCT

(10) International Publication Number  
WO 00/33160 A3

- (51) International Patent Classification<sup>7</sup>: H04N 7/173
- (21) International Application Number: PCT/US99/28528
- (22) International Filing Date: 2 December 1999 (02.12.1999)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
09/204,888 3 December 1998 (03.12.1998) US  
09/205,119 3 December 1998 (03.12.1998) US  
09/205,653 3 December 1998 (03.12.1998) US
- (81) Designated States (*national*): AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, VN, YU, ZA, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- (71) Applicant (*for all designated States except US*): TELECOM PARTNERS LTD. [US/US]; 300 N. Broad Street, Doylestown, PA 18901 (US).
- Published:  
— With international search report.
- (88) Date of publication of the international search report:  
30 November 2000
- (72) Inventors; and
- (75) Inventors/Applicants (*for US only*): ELDERING, Charles, A. [US/US]; 315 Hedgerow Lane, Doylestown, PA 18901 (US). SYLLA, M., Lamine [SN/US]; 6 West Butler Avenue, New Britain, PA 18901 (US).
- (74) Agent: RYDER, Douglas, J.; 300 North Broad Street, Doylestown, PA 18901 (US).

(54) Title: SUBSCRIBER CHARACTERIZATION AND ADVERTISEMENT MONITORING SYSTEM



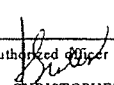
(57) Abstract: A subscriber characterization and advertisement monitoring system is presented in which subscriber viewing habits are monitored to determine demographic profiles. These profiles can be utilized for the matching of advertisements to subscribers based on their viewing habits and estimated demographics and product interests. The system can be run locally in a television set-top (1808) or can be run in client server mode where channel selections are transmitted from the residence (1800) to a centralized switching location server (1840) such as a telephone office or Internet Service Provider. In client-server mode the channel selections are monitored at the centralized location (1840) which also performs the subscriber characterization. The system also provides the ability to monitor if advertisements were viewed and for what duration.

WO 00/33160 A3



## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US99/28528

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> IPC(7) : H04N 7/173 US CL : 345/327; 348/1,7,9,12,13; 455/2,4,2, 5,1,6,2,6,3 According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) U.S. : 345/327; 348/1,7,9,12,13; 455/2,4,2, 5,1,6,2,6,3  Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US, 5,374,951 A (WELSH) 20 December 1994 figure 7 & columns 14-15.	1
A, P	US 5,977,964 A (WILLIAMS et al.) 02 November 1999 whole document	1-7
A	US 5,786,845 A (TSURIA) 28 July 1998 whole document	8-13
A	US 5,233,423 A (JERNIGAN et al.) 03 August 1993 whole document	8-13
A	US 5,805,974 A (HITE et al.) 08 September 1998 whole document	8-13
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* "A"	Special categories of cited documents: document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"B"	earlier document published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document referring to an oral disclosure, use, exhibition or other means	"A" document member of the same patent family
"P"	document published prior to the international filing date but later than the priority date claimed	
Date of the actual completion of the international search 16 JUNE 2000		Date of mailing of the international search report 01 AUG 2000
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230		Authorized officer  CHRISTOPHER GRANT Telephone No. (703) 305-4755

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US99/28528

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X -- Y	US 5,758,259 A (LAWLER) 26 May 1998 figures 5 & 6.	14,15,22 33,34,40 ----- 30-32, 49-51
X	US 5,635,989 A (ROTHMULLER) 03 June 1997 figure 3	14,22
X	US 5,410,344 A (GRAVES et al.) 25 April 1995 whole document, especially figure 4.	14,16-18, 22,23
A,E ---- Y,E	US 6,002,394 A (SCHEIN et al.) 14 December 1999 col. 15, line 58 - col. 16, line 14.	25-29,44-48 ----- 30-32,49-51
A	US 5,515,098 A (CARLES) 07 May 1996 whole document	30-32,49-51